

TARTU ÜLIKOOL  
LOODUS- JA TEHNOLOOGIATEADUSKOND  
MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT  
BIOTEHNOLOOGIA ÕPPETOOL

Anett Tähiste

**Geneetilistel markeritel põhinev südame isheemiatõve riski  
hindamine Eesti Geenivaramu kohordis**

Bakalaureusetöö

Juhendaja MSc Maris Alver  
Kaasjuhendajad MSc Kristi Läll  
MD, PhD Andres Metspalu

TARTU 2014

# SISUKORD

KASUTATUD LÜHENDID .....	3
SISSEJUHATUS .....	4
1. KIRJANDUSE ÜLEVAADE.....	5
1.1 Üldiseid aspekte.....	5
1.2 Ateroskleroosi molekulaarsed alused .....	6
1.3 Südame isheemiatõve geneetilised uuringud.....	8
1.3.1 Aheldusuuringud .....	8
1.3.2 Assotsiatsiooniuuringud .....	9
1.3.2.1 Kandidaatgeenide uuringud.....	10
1.3.2.2 Ülegenoomsed assotsiatsiooniuuringud .....	11
1.4 Riskiskoorid.....	13
1.4.1 Framinghami riskiskoor.....	14
1.4.2 Muud riskiskoorid.....	16
1.4.3 Geneetiline riskiskoor.....	17
2. EKSPERIMENTAALOSA .....	19
2.1 Töö eesmärk .....	19
2.2 Materjalid ja meetoodika.....	19
2.2.1 Valim .....	19
2.2.2 Geneetiliste markerite valik.....	20
2.2.3 Andmete analüüsitavale kujule viimine .....	21
2.2.4 Statistiline analüüs ja riskiskoori arvutamine .....	22
2.3 Tulemused .....	22
2.4 Arutelu .....	24
KOKKUVÕTE .....	28
SUMMARY.....	29
TÄNUSÕNAD .....	31
KIRJANDUSE LOETELU.....	32
LISA .....	41
LIHTLITSENTS.....	43

## KASUTATUD LÜHENDID

CAD	südame isheemiatõbi ( <i>coronary artery disease</i> )
GWAS	ülegenoomne assotsiatsiooniuuring ( <i>genome-wide association study</i> )
HLA	inimese leukotsüüdi antigeen ( <i>human leukocyte antigen</i> )
ICAM-1	intratsellulaarne adhesioonimolekul 1 ( <i>intracellular adhesion molecule-1</i> )
LD	ahelduse tasakaalutus ( <i>linkage disequilibrium</i> )
LDL	madala tihedusega lipoproteiin ( <i>low-density lipoprotein</i> )
Lp-PLA2	lipoproteiiniga seostunud fosfolipaas A2 ( <i>lipoprotein-associated phospholipase A2</i> )
Lp(a)	lipoproteiin(a) ( <i>lipoprotein(a)</i> )
SMC	silelihasrakk ( <i>smooth muscle cell</i> )
SNP	ühenukleotiidne polümorfism ( <i>single nucleotide polymorphism</i> )
VLDL	väga madala tihedusega lipoproteiin ( <i>very low density lipoprotein</i> )

## SISSEJUHATUS

Kardiovaskulaarhaigused põhjustavad ligi 1/3 surmajuhtumitest maailmas ning haigestumine nendesse tõvedesse näib omavat pidevat kasvutendentsi. Kuigi viimased 20 aastat on tehtud märkimisväärsed edusamme südame isheemiatõve varaseks avastamiseks ja eduka ravimeetodi väljatöötamiseks, on see jätkuvalt peamine surma põhjustav haigus enamikus maailmariikides. Epidemioloogiliste ja populatsioonipõhiste uuringute tulemusena on identifitseeritud peamised südame isheemiatõve arengu eest vastutavad riskifaktorid. Kuna mitmetel südame-veresoonkonnahaigusi põdevatel indiviididel puuduvad tavapärased riskifaktorid, ei võimalda ainuüksi nende tegurite hindamine kardiovaskulaarse haiguse tekke riski efektiivselt ennustada. Molekulaarsete ja kliiniliste katsete ning kandidaatgeenidel põhinevate teadustööde tulemusena on täheldatud positiivseid leide südame isheemiatõvega seotud geenide kohta, kuid haiguse kompleksuse tõttu on patogeneesi detailne mõistmine siiani raskendatud. Tänu biotehnoloogia ja -informaatika hüppelisele arengule on viimasel kümnendil hakatud südame isheemiatõve tekke ja kuluga seotud geneetiliste variantide tuvastamisel rakendada ülegenoomseid assotsiatsiooniuuringuid. Geneetiliste markerite uurimine võib aidata indiviidide haigestumisriski täpsemalt hinnata.

Käesoleva töö eesmärgiks oli hinnata geneetiliste markerite võimet ennustada südame isheemiatõve esinemist Eesti Geenivaramu valimis.

# 1. KIRJANDUSE ÜLEVAADE

## 1.1 Üldiseid aspekte

Südame-veresoonkonnahaigused ja nendega kaasnevad tüsistused on peamised surma põhjustavad haigused arenenud riikides ning ennustatavalt aastaks 2020 ka arengumaades (Dahlöf, 2010; Celermajer jt., 2012). Igal aastal sureb Euroopas kardiovaskulaarhaiguste, näiteks südame isheemiatõve (*coronary artery disease* – CAD) või ajuinsuldi tagajärjel üle nelja miljoni inimese (Nichols jt., 2012). Eesti on südame-veresoonkonnahaigustesse suremuse poolest maailmas esimeste hulgas<sup>1</sup>. Südame isheemiatõbi, mis tekib geneetiliste ja keskkondlike tegurite ning nende omavaheliste interaktsioonide tulemusena, on keeruline põletikuline haigus, mida iseloomustab südant hapnikuga varustavate koronaar- ehk pärgarterite ahenemine (Wang, 2005). CAD-i ja selle erinevate kliiniliste ilmingute (näiteks müokardiinfarkt või rinnaangiin) peamine põhjus on ateroskleroos (Libby ja Theroux, 2005; Cassar jt., 2009).

Peamised CAD-i tekke ja kulu eest vastutavad riskifaktorid on meessugu, vanus, vere suurenenud kolesteroolisisaldus ehk hüperkolesteroleemia, diabeet, kõrge vererõhk, ülekaalulisus, suitsetamine ja vähene füüsiline aktiivsus (Swerdlow jt., 2012). Nimetatud tegurite hindamisel pole indiviidi kardiovaskulaarse haiguse ja tüsistuste tekke riski võimalik aga täielikult ennustada. Näiteks 35% CAD-i juhtudest esineb indiviididel, kellel üldkolesteroolitase on normi piires (<200 mg/dL) (Castelli, 1996). Lisaks umbes 20% südame-veresoonkonnahaigusi põdevatest patsientidest ei oma ühtki loetletud riskifaktorit (Khot jt., 2003; Hozawa jt., 2007). Seetõttu on viimastel aastatel suurt tähelepanu osutatud uudsete riskifaktorite, sealhulgas geneetiliste markerite uurimisele. Kuigi mitmed teadustööd on kinnitanud, et südame isheemiatõbi on suures osas päritav, ei ole pärilikkuse molekulaarsed mehhanismid tõve kompleksuse tõttu veel täpselt teada (Baudhuin, 2009; Peden ja Farrall, 2011).

Südame isheemiatõve geneetiliste aluste selgitamine on haigusele keskenduvate biomeditsiiniliste uuringute üheks põhiprioriteediks. Uute molekulaargeneetiliste katsedisainide väljatöötamine, haigusseoseliste ühenukleotiidsete polümorfismide (*single*

---

<sup>1</sup> [http://www.who.int/cardiovascular\\_diseases/publications/atlas\\_cvd/en/](http://www.who.int/cardiovascular_diseases/publications/atlas_cvd/en/)

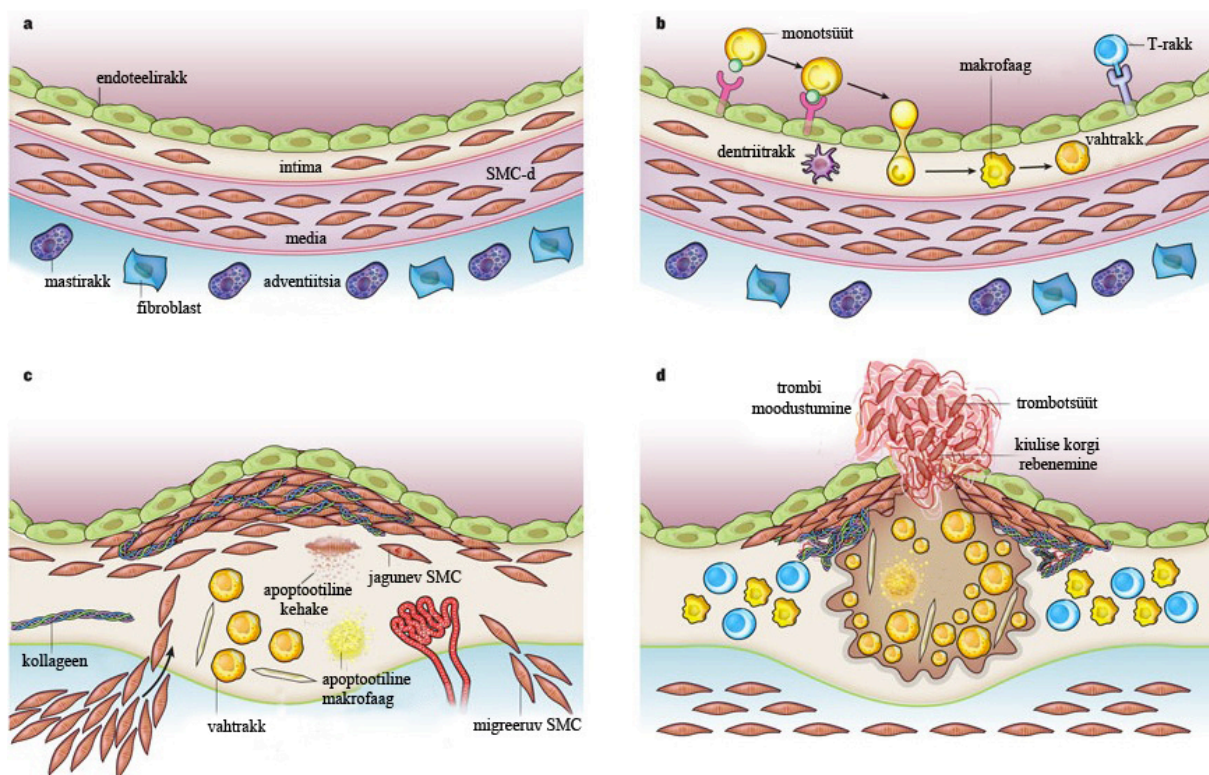
*nucleotide polymorphism* – SNP) kaardistamine inimese genoomis ja bioinformaatiliste meetodite areng uute haigust soodustavate geenivariantide identifitseerimiseks on suureks panuseks täpsema riski hindamise ning haiguse varasema diagnoosimise või täieliku vältimise kriteeriumide väljatöötamisel. Epidemioloogiliste uuringute ja kliiniliste katsete tulemusena on esitatud veenvaid tõendeid selle kohta, et südame-veresoonkonnahaigused on suures osas ennetatavad (Cooper jt., 2000).

## **1.2 Ateroskleroosi molekulaarsed alused**

Ateroskleroosist tingitud tüsistused on peamised surmapõhjustajad lääneriikides (Glass ja Witztum, 2001). Ateroskleroos, südame isheemiatõve peamine põhjus, on multifaktoriaalne põletikuline haigus, millele on iseloomulikud suurte ja keskmiste arterite sisekestale ladestunud aterosklerootilised naastud, mis võivad põhjustada arterite ahenemist, kõvastumist ja/või täielikku ummistust (Greaves ja Channon, 2002). Kui kõrge madala tihedusega lipoproteiinide (*low-density lipoprotein* – LDL) taset veres on peetud peamiseks ateroskleroosi riskiteguriks, siis viimase 20 aasta jooksul on märkimisväärselt suurenenud huvi haiguse põletikuliste mehhanismide vastu (Ross, 1999). Kuigi kliinilised kardiovaskulaarsed juhud (näiteks müokardiinfarkt või ajuinsult) ilmnevad tavaliselt meestel pärast viiekümnendat eluaastat ja naistel pärast kuuekümnendat eluaastat, siis tegelikult algab ateroskleroosi kujunemine juba varases eas, isegi lootelises arengustaadiumis (Mecchia jt., 2009; Maas ja Appelman, 2010).

Ateroskleroos on progressiivne krooniline protsess, mis saab alguse erinevate riskitegurite (näiteks hüperkolesteroleemia, kõrge vererõhk, suitsetamine) poolt tingitud endoteeli vigastusest ja sellega kaasnevast põletikulisest reaktsioonist (Chhabra, 2009). Kahjustunud endoteel ja muutused ekstratsellulaarse maatriksi koostises võimaldavad LDL-partiklite sisenemist arterite sisepinda vooderdavasse rakukihti ehk intimasse (Tabas jt., 2007). Seal leiab aset nende modifitseerimine erinevate oksüdeerijate (näiteks müeloperoksüdaas, vabad radikaalid) ja/või proteolüütiliste, lipolüütiliste ning hüdrolyütiliste ensüümide toime tõttu (Öörni jt., 2000). Oksüdeeritud madala tihedusega lipoproteiine peetakse kõige aterogeensemateks LDL-partikliteks (Catapano jt., 2000). Aktiveeritud endoteelirakud eritavad kemokiine, sealhulgas monotsüütide kemotaktilist valku 1 (MCP-1) ja interleukiini (IL)-8, leukotsüütide adhesioonimolekule, nagu näiteks intratsellulaarset adhesioonimolekuli 1 (*intracellular adhesion molecule-1* – ICAM-1), vaskulaarset adhesioonimolekuli 1 (VCAM-

1) ning E- ja P-selektiine, mis indutseerivad monotsüütide, lümfotsüütide ja polümorfonukleaarsete granulotsüütide sissetungi intimasse (Businaro jt., 2012). Monotsüütidest diferentseerunud makrofaagid seonduvad oksüdeeritud LDL-partiklitega, püüdes neid fagotsüteerides kõrvaldada ning muutuvad selle tulemusena kolesterooli täisvalgunud vahtrakkudeks (Glass ja Witztum, 2001; Ghattas jt., 2013). Moodustub esimene aterosklerootiline kahjustus – rasvatriip, millega ei kaasne küll haigussümptomeid, kuid mis võib edasi areneda aterosklerootiliseks naastuks (Hansson, 2005). Pidev rakulise ja humoraalse immuunvastuse aktiveerimine viib kroonilise põletikulise seisundini (Libby, 2012; Witztum ja Lichtman, 2014). Protsess jätkub silelihasrakkude (*smooth muscle cell* – SMC) migratsiooniga arteri mediaalsest kihist intimasse, millega kaasneb komplekssema aterosklerootilise kahjustuse teke (Glass ja Witztum, 2001). Intimas olevad SMC-d toodavad ekstratsellulaarse maatriksi molekule (näiteks kollageen ja elastiin), luues naastu katva kiulise korgi. Kiulise korgi all olevas naastus läbivad vahtrakud apoptoosi, vabastades oma lipiidse sisu, mille tulemusena moodustub lipiidirikas nekrootiline südamik (Tabas, 2010). Kiulise korgi õhenemine vähendab aterosklerootiliste naastude stabiilsust, muutes need vastuvõtlikumaks rebenemisele. Korgi purunemisel võib naastu sisu vabaneda vereringesse ja põhjustada trombi moodustumist (Moore ja Tabas, 2011). Kui tekkinud tromb takistab verevoolu südamesse, on tulemuseks südameinfarkt, kui aga on blokeeritud vere juurdevool ajju, siis on tagajärjeks ajuinsult (joonis 1) (Willyard, 2013).



**Joonis 1.** Normaalse lihaselise arteri ja rakkude muutused ateroskleroosi progresseerumisel

tromboosi tekkeni. a: Arteri sein moodustavad kolm kihti: endoteelirakkudega vooderdatud silelihasrakkudest koosnev sisekiht ehk intima, ekstratsellulaarse matriksiga kompleksis olevatest silelihasrakkudest koosnev keskkiht ehk media ning mastirakkudest, sidekoerakkudest, närvilõpmetest ja mikroveresoontest koosnev väliskiht ehk adventiitsia. b: Ateroskleroosi esmases etapis toimub leukotsüütide adhesioon aktiveeritud endoteelirakkudele ja sisenemine intimasse ning monotsüütide diferentseerumine makrofaagideks, mis omakorda muutuvad vahtrakkudeks. c: Kahjustuse progresseerumine hõlmab medias olevate silelihasrakkude migratsiooni intimasse, nende proliferatsiooni ning ekstratsellulaarse matriksi molekulide kõrgeks kogunenud sünteesimist. Vahtrakud läbivad apoptoosi, mille tulemusena neist vabanenud ekstratsellulaarsed lipiidid akumulereuvad ateroskleroosilise naastu keskele, moodustades nekrootilise südamiku. d: Ateroskleroosilist naastu katva kiulise korgi rebenemisel võib veresoone luumenini ulatuv tromb takistada verevoolu (Libby jt., 2011, kohandatud). SMC – silelihasrakk.

### **1.3 Südame isheemiatõve geneetilised uuringud**

Sõltuvalt pärilikkuse muustritest saab haigused klassifitseerida kahte suurde kategooriasse: monogeensed haigused, mida põhjustab üks kindel defektne geen, ja kompleksed haigused, mida iseloomustavad mitmete geenide ja keskkonnategurite omavahelised interaktsioonid (Chial, 2008; Lobo, 2008). Kuna paljude monogeensete haiguste (näiteks perekondlik hüperkolesteroleemia) geneetiline taust on tänaseks teada, on biomeditsiiniliste uuringute üheks praeguseks põhieesmärgiks inimese genoomi uurides tuvastada komplekssete haiguste kujunemise eest vastutavaid spetsiifilisi geene ja DNA järjestuse variante (Civeira jt., 2004; Kathiresan ja Srivastava, 2012). Geneetilise pildi selgitamine lubaks paremini mõista vastuvõtlikkust nendele haigustele. Uute haigusrisi suurendavate geneetiliste variatsioonide tuvastamine võimaldab juhtida tähelepanu seni tundmatutele bioloogilistele radadele ja seeläbi täiustada diagnoosi- ja ravimeetodeid (Musunuru ja Kathiresan, 2010). Haiguse geneetiliste aluste uurimiseks rakendatakse peamiselt aheldus- ja assotsiatsiooniuuringuid (Elosua jt., 2009).

#### **1.3.1 Aheldusuuringud**

Aheldusuuringute eesmärk on kindlaks määrata haigusele vastuvõtlikkust soodustavad või seda põhjustavad genoomipiirkonnad (Sayols-Baixeras jt., 2014). Katsedisain põhineb huvipakkuva fenotüübiga perekonnaliikmete päriliku materjali võrdlemisel üle mitme põlvkonna (Abbate jt., 2008). Selleks vaadatakse ühtlaselt üle kogu genoomi jaotunud geneetilisi markereid (näiteks mikrosatelliidid) ning analüüsitakse haigusseoseliste



kromosoomipiirkondade pärandumist põlvest põlve. Huvipakkuva regiooni kaardistamine lubab tuvastada põhjusliku geeni (Kathiresan ja Srivastava, 2012).

Aheldusuuring on väga hea katsedisain suhteliselt harva esinevate monogeensete haiguste uurimisel (Pericak-Vance, 2001). Selle uuringutüübi eeliseks on fenotüübiliselt hästi iseloomustatud perekondade ja informatiivsete indiviidide kaasamine (Abbate jt., 2008). Sellest lähtuvalt on seda katsedisaini edukalt rakendatud mendellike südame-veresoonkonnahaiguste, näiteks perekondliku hüperkolesteroleemia uurimisel (Dietter jt., 2004). Tegemist on lipoproteiinide metabolismi häirega, millele on iseloomulik kõrge LDL-kolesterooli tase veres, varane müokardiinfarkti esinemine ja ksantoomide ehk lipiidikogumike ladestumine nahal (Marks jt., 2003). Enamik haigusjuhte on põhjustatud mutatsioonidest LDL-retseptorit kodeerivas geenis (*LDLR*), kuid haigust põhjustavaid geneetilisi muutusi on täheldatud ka LDL-retseptori ligandina funktsioneeriva apolipoproteiini geenis (*APOB*) ja LDL-retseptori lagundamisel osaleva valgu geenis (*PCSK9*) (Rader jt., 2003; Damgaard jt., 2004). Ülegenoomse aheldusuuringuga on leitud veel seos müokardiinfarkti ja *ALOX5AP* geeni (Helgadottir jt., 2004), südame isheemiatõve ja *MEF2A* geeni (Wang jt., 2003) ning kolesterooli metabolismi ja *PCSK9* geeni (Varret jt., 1999; Chen jt., 2005) vahel. Lisaks on analüüsitud ka ateroskleroosi kvantitatiivseid markereid, näiteks pärgarteri kaltsiumitaset (Lange jt., 2002), unearteri intima-media paksust (Fox jt., 2004) ja pahkluu-õlavarre indeksit (Kullo jt., 2006), kuid kandidaatgeene ei ole veel tuvastatud.

### 1.3.2 Assotsiatsiooniuuringud

Assotsiatsiooniuuringud on geneetiliselt komplekssete haiguste uurimiseks efektiivsemad kui aheldusuuringud, sest neile on omane suurem statistiline võimsus avastada väikese efektiga geneetilisi variante (Risch ja Merikangas, 1996). Nende uuringutega hinnatakse geneetiliste variantide (tavaliselt SNP-de või koopiarvu variatsioonide) ja haiguse olemasolu/puudumise või konkreetse fenotüübi väärtuse vahelist seost (Sayols-Baixeras jt., 2014). Kõige sagedamini kasutatakse assotsiatsiooniuuringutes juht-kontroll tüüpi katsedisaini, kus mittesuguluses olevate tervete (kontrollid) ja haigust põdevate indiviidide (juhud) vaheline alleelisageduste võrdlus lubab hinnata geneetiliste variantide mõju fenotübile (Kruglyak, 2005).

Assotsiatsiooniuuringute peamine eelis on nende lihtsus ja paindlikkus, võimaldades uurida geen-geen ja geen-keskkond vahelisi interaktsioone. Samas võivad mitmed puudused mõjutada tulemuste usaldusväärsust: täpse fenotüübi defineerimise keerukus (Stranger jt., 2011), väikese valimi suurusest tulenevad valepositiivsed tulemused (Roberts ja Stewart, 2012a) ja tulemuste mittereplitseeritavus (Gorroochurn jt., 2007). Väikese efektiga geneetiliste variantide identifitseerimine nõuab suurt valimit, mis saavutatakse tavaliselt erinevate valimite ja populatsioonide liitmisel. Paljudel juhtudel kaasneb heterogeense valimiga aga populatsioonispetsiifiliste variantide leidmise võimsuse vähenemine (Sayols-Baixeras jt., 2014).

Laialdaselt kasutatud leidnud assotsiatsioonanalüüsid on kandidaatgeenide uuringud ja ülegenoomsed assotsiatsiooniuuringud (*genome-wide association study* – GWAS).

### **1.3.2.1 Kandidaatgeenide uuringud**

Kandidaatgeenide assotsiatsiooniuuringuid on laialdaselt rakendatud komplekssete haiguste analüüsimisel (Zhu ja Zhao, 2007). Nende uuringute eesmärk on hinnata konkreetse alleeli või geenivariandi mõju haiguse kujunemisele. Selle meetodi puhul valitakse kindlad kandidaatgeeni polümorfismid ja testitakse nende sagedust ja ahelduse tasakaalutust (*linkage disequilibrium* – LD) juhuslikes valimites, mis koosnevad nii tervetest kui ka haigust põdevatest indiviididest (Kwon ja Goate, 2000). Kandidaatgeeniuuringute tulemusena on tuvastatud enam kui 300 südame isheemiatõvega seotud geeni, mis funktsioneerivad paljudes erinevates molekulaarsetes radades, näiteks lipiidide metabolismis, vere hüübimises, vererõhu, põletiku ja rakutsükli regulatsioonis (Pranavchand ja Reddy, 2013). Vaatamata arvukatele leidudele on väheseid tulemusi aga replitseeritud. Üheks põhjuseks peetakse ebapiisavat valimi suurust (Zeller jt., 2012).

Üheks positiivseks kandidaatgeeniuuringu tulemuseks on enam kui 50 aastat tagasi täheldatud seos ABO lookuse ja müokardiinfarkti vahel (Bronte-Stewart jt., 1962). Tänapäevaks on vastavat seost kinnitanud ka GWAS-i tulemused (Reilly jt., 2011). Molekulaarsed katsed on näidanud, et A, B ja O on sama geeni erinevad alleelid. A ja B alleelilt kodeeritakse transferaasi, mis lisab karbohüdraadi molekuli von Willebrandi faktorile, pikendades selle pooleluiga. See suurendab omakorda tromboosi ja müokardiinfarkti riski. O alleeliga valgul puudub

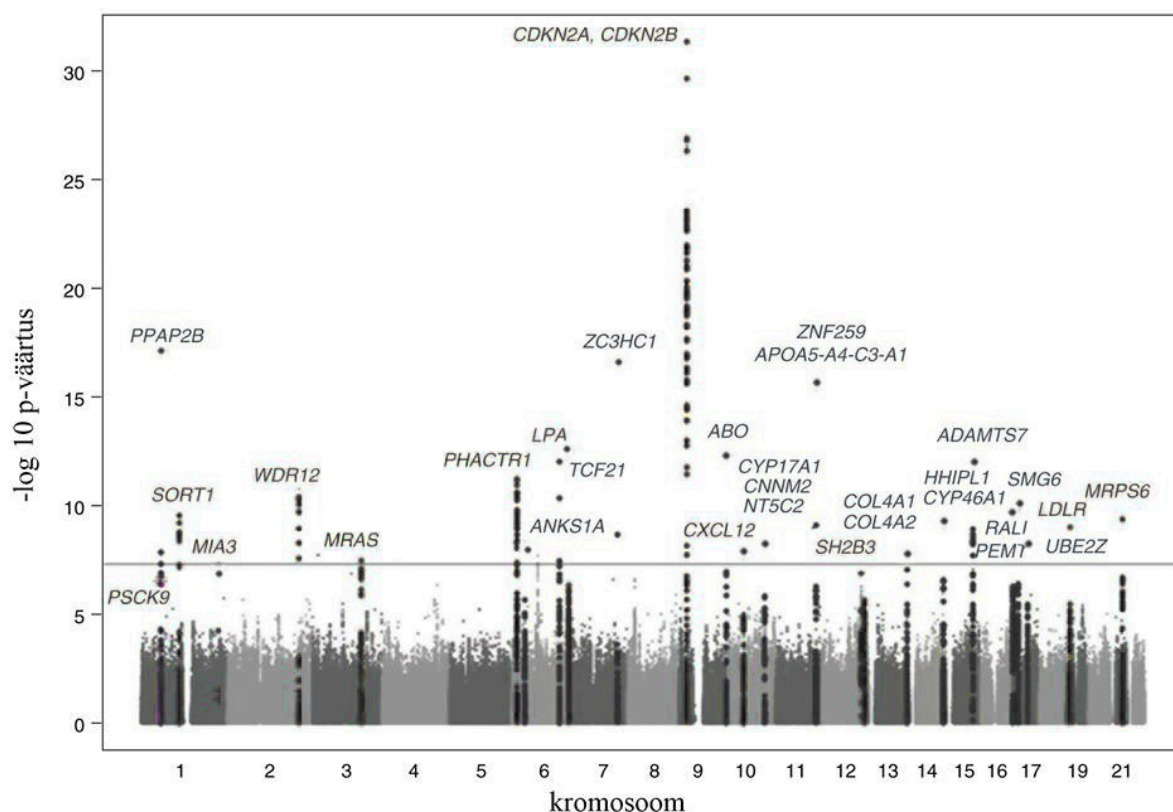
punktmutatsiooni tõttu aga ensümaatiline aktiivsus, mistõttu pole O veregruppi kandvatel indiviididel ka kõrgenenud riski (Roberts ja Stewart, 2012b). Samuti on kandidaatgeeniuuringute tulemusena selgunud, et koopiaarvu variatsioonid *LPA* geenis mõjutavad lipoproteiin(a) [Lp(a)] taset veres (Lusis jt., 2004) ja on seotud müokardiinfarkti riskiga (Kamstrup jt., 2009). Lp(a), mis koosneb apolipoproteiin(a)-ga kovalentselt seotud kolesteroolirikkast LDL-partiklist, omab olulist rolli ateroskleroosi tekkes (Trégouët jt., 2009), soodustades põletikulist vastust (Poon jt., 1997), vahtrakkude moodustumist (Nielsen jt., 1998) ning tromboosi teket (Boffa jt., 2004).

### 1.3.2.2 Ülegenoomsed assotsiatsiooniuuringud

Kui aheldusuuringud on olnud edukad monogeensete haiguste harvade variantide tuvastamisel, siis ülegenoomsete assotsiatsiooniuuringute katsedisain võimaldab avastada levinud haiguste sagedasi variante (Musunuru ja Kathiresan, 2010). GWAS-i eesmärgiks on tuvastada komplekssete fenotüüpidega seotud geneetilisi variante eelneva kandidaatgeenide või -lookuste defineerimiseta (Visscher jt., 2012). GWAS sai võimalikuks tänu tehnoloogilistele arengutele, mis võimaldasid suure hulga SNP-de genotüpiseerimist ja ahelduse tasakaalutuse mustrite määramist erinevates populatsioonides (International HapMap Consortium jt., 2007). GWAS-i katsedisainiga kaasnevate probleemide ja piirangute tõttu peab nende uuringute läbiviimine olema hoolikalt teostatud. Väga oluline on hoida juhtude fenotüübiline heterogeensus ja kontrollide väärklassifikatsioon miinimumi lähedal (McCarthy jt., 2008). Lisaks on oluline, et valimi suurus ulatuks tuhandete, enamasti isegi sadade tuhandete indiviidideni. GWAS-i puhul on risk valepositiivseteks tulemusteks suur, mistõttu on range kvaliteedikontroll ning tulemuste replitseerimine sõltumatu valimil hädavajalik (Baudhuin, 2009; Swerdlow jt., 2012).

Ülegenoomsetes assotsiatsiooniuuringutes on kaardistatud üle 40 südame isheemiatõvega seotud geneetilise regiooni (Sayols-Baixeras jt., 2014). Enamasti mittekodeerivas piirkonnas paiknevate SNP-de hulgas on nii riskifaktoritega seotud kui ka mitteseotud markereid. Seni läbiviidud GWAS-ide tulemusena on leitud tugevaim seos südame isheemiatõve ja kromosoomi 9p21.3 lookuses olevate SNP-de vahel (joonis 2) (Roberts ja Stewart, 2012b). Vastav lookus tuvastati 2007. aastal kolmes erinevas samaaegselt läbiviidud uuringus (Helgadottir jt., 2007; McPherson jt., 2007; Wellcome Trust Case Control Consortium, 2007).

Assotsieerunud markerid, mis on omavahel tugevasti aheldunud, paiknevad geenivaeses regioonis, mis sisaldab 126 000 aluspaari pikkust mittekodeerivat RNA-d *CDKN2B-AS1* (*CDKN2B antisense RNA 1*) (Roberts ja Stewart, 2012b). Lähimad valke kodeerivad geenid paiknevad 100 kiloaluspaari kaugusel. Nendeks on *CDKN2A* ja *CDKN2B*, millelt kodeeritakse rakkude proliferatsiooni ja apoptoosi kontrollimisel osalevaid rakulise vananemise inhibiitoreid (Swerdlow jt., 2012). On leitud, et ühe koopiaga indiviididel on 15-20% ja kahe koopia kandjatel 40% võrra suurem CAD-i risk, ning riskihaplotüüpi kannab 75% Euroopa elanikkonnast (McPherson jt., 2007; Roberts ja Stewart, 2012b). Vastava lookuse ja CAD-i kujunemise riski vahelist seost on näidatud ka teistel rahvustel, sealhulgas hiinlastel (Ding jt., 2009), korealastel (Shen jt., 2008), jaapanlastel (Hinohara jt., 2008), indialastel (Kumar jt., 2011) ning pakistanlastel (Saleheen jt., 2010). Kromosoomi 9p21.3 lookust on seostatud veel teiste fenotüüpidega, näiteks teist tüüpi diabeedi (Shea jt., 2011) ning kõhuaordi ja aju aneurüsmiga (Helgadottir jt., 2008), kuid klassikaliste CAD-i riskifaktoritega seosed puuduvad (Chen jt., 2009). Uuringud on geneetiliste variantide ja suurenenud CAD-i riski analüüsimisel viidanud mitmetele potentsiaalsetele molekulaarsetele mehhanismidele, näiteks *CDKN2B-AS1* ekspressiooni muutusele ja põletiku regulatsioonis osaleva transkriptsioonifaktori STAT1 seondumiskoha kadumisele, kuid selget põhjust pole veel leitud (Kathiresan ja Srivastava, 2012; Zeller jt., 2012).



**Joonis 2.** Manhattani joonis CAD-i GWAS-i tulemustest >100 000 eurooplase seas. X-teljel

on toodud kromosoominumber ning y-teljel p-väärtuste kümnendlogaritm. Geeninimedega on tähistatud kõige tõenäolisemad kandidaatgeenid CAD-i põhistes GWAS-i töödes identifitseeritud lookustes. Horisontaalne joon märgib GWAS-i tööde statistilise olulisuse nivood ( $p\text{-väärtus} < 5 \times 10^{-8}$ ) (Schunkert jt., 2011, kohandatud).

Teine tähelepanuväärne GWAS-i leid on 2007. aastal identifitseeritud lookus 1p13.3, mida on mitmetes uuringutes seostatud ka LDL-kolesterooli tasemega (Samani jt., 2007; Muendlein jt., 2009; Teslovich jt., 2010). Selles lookuses paikneb kolm geeni: *CELSR2*, *PSRC1* ja *SORT1*, millest huvipakkuvaim on viimane. *SORT1* poolt kodeeritav valk sortilin, mida leidub mitmetes intratsellulaarsetes kompartmentides (näiteks Golgi kompleks), osaleb nii endotsütoosis kui ka valkude rakusiseses transpordis (Nielsen jt., 2001). Funktsionaalsed katsed näitasid, et vastava lookuse mittekodeerivas regioonis asuva SNP-i rs12740374 minoorne alleel loob transkriptsioonifaktorile C/EBP $\alpha$  (CCAAT-võimendajaga seonduv valk) seondumiskoha, millega kaasneb kõrgeenenud *SORT1* ekspressioonitase maksarakkudes. Uuringus selgus, et *SORT1* üleekspressioon hiirtes vähendas väga madala tihedusega lipoproteiinide (*very low density lipoprotein* – VLDL) sekretsiooni maksarakkudest, alandades seeläbi LDL-kolesterooli taset veres (Musunuru jt., 2010). Teises uuringus demonstreeriti aga vastupidist efekti: madala *SORT1* ekspressioonitasemega kaasnes vähenenud VLDL-partiklite sekretsioon ning LDL-kolesterooli taseme langus veres (Kjolby jt., 2010). Arvatakse, et vastupidised tulemused on seotud erinevate katsetingimuste ja *knockout* mudelitega (Tall ja Ai, 2011).

## 1.4 Riskiskoorid

Südame-veresoonkonnahaigustesse suremuse suure sageduse tõttu on kliinilise kardioloogia suurimaks väljakutseks tuvastada haigestumisriskiga patsiendid. Südame isheemiatõve ilmnemist ja kulgu ei ole võimalik küll täpselt ennustada, kuid erinevatel kohortuuringutel põhinevate mudelite kasutamine lubab haiguse ja tüsistuste tekke riski siiski hinnata (Dent, 2010a). Välja on arendatud mitmeid riski arvutamise mudeleid, mis demograafilisi ja kliinilisi tunnuseid arvesse võttes võimaldavad patsiente jagada madala, keskmise ja kõrge riskiga kategooriasse (Assmann jt., 2002; Nucifora jt., 2009).

Riskiskoori arvutamise mudel peaks sisaldama haiguse riskifaktoreid (vanus, sugu, lipiidide tase, süstoolse vererõhu väärtus, suitsetamisharjumused, perekondlik eelsoodumus) ning olema täpne antud populatsiooni jaoks (McGinn jt., 2000; Beswick ja Brindle, 2006).

Patsiendi riskiastme kindlakstegemine lubab seejärel määrata sobivaima ravimeetodi: iniviidid, kellel on diagnoositud südame-veresoonkonnahaigus (näiteks perifeersete arterite haigus või aordianeurüsm) kategoriseeritakse automaatselt kõrge kardiovaskulaarse riski rühma ja ravi põhineb sekundaarse ennetamise strateegiatel; tervete patsientide puhul kasutatakse primaarset ennetamist, mille puhul kõrge riskiga iniviididel rakendatakse kohe medikamentooset ravi (näiteks lipiidide taset alandavad preparaadid), keskmisesse riskikategooriasse kuuluvatele patsientidele määratakse samuti ravimid või jätkatakse mitteinvasiivset testimist ning madala riskiga iniviididele soovitatakse vajadusel elustiili muutmist (Lloyd-Jones, 2010).

#### **1.4.1 Framinghami riskiskoor**

Enim kasutatud mudel on Framinghami riskiskoor, mis traditsiooniliste riskitegurite (vanus, sugu, süstoolne vererõhk, üld- ja HDL-kolesterooli tase, suitsetamisharjumused) põhjal võimaldab hinnata iniviidi kardiovaskulaarse haigusjuhu tekke riski järgmise kümne aasta jooksul (Wilson jt., 1998). Kuigi Framinghami mudelit peetakse oluliseks vahendiks ennetavas kardioloogias, on selle rakendamisel täheldatud mitmeid kitsaskohti (Davis jt., 2010).

Oluline probleem Framinghami mudeli kasutamisel on selle üldistatavuse puudumine. Kuna Framinghami südameuuringusse, mis viidi läbi 1948. aastal peamiste kardiovaskulaarse riski tegurite kirjeldamiseks, kaasati enamasti ainult valgeid keskklassi kuuluvaid ameeriklasi, siis ei pruugi selle mudeli tulemused teiste populatsioonide ja etniliste rühmade puhul täpsed olla (Davis jt., 2010). D'Agostino ja tema kolleegide poolt läbiviidud uuringus selgus, et Framinghami mudel hindab riski jaapani päritoluga ameeriklaste, hispaaniakeelsete ameeriklaste (*Hispanic Americans*) ja põlisameerika naiste puhul üle (D'Agostino jt., 2001). Teises uuringus täheldati, et nimetatud mudel alahindab riski aasialaste hulgas (Kanjilal jt., 2008). Seetõttu on loodud Framinghami mudelil põhinev riskikalkulaator, mis arvestab ka patsiendi etnilist kuuluvust<sup>2</sup>.

Enamik riskimudeleid on välja töötatud keskealiste seas, näiteks Framinghami riskimudel loodi iniviidide põhjal, kelle keskmine vanus oli 49 aastat (Wilson jt., 1998). Kuna on

---

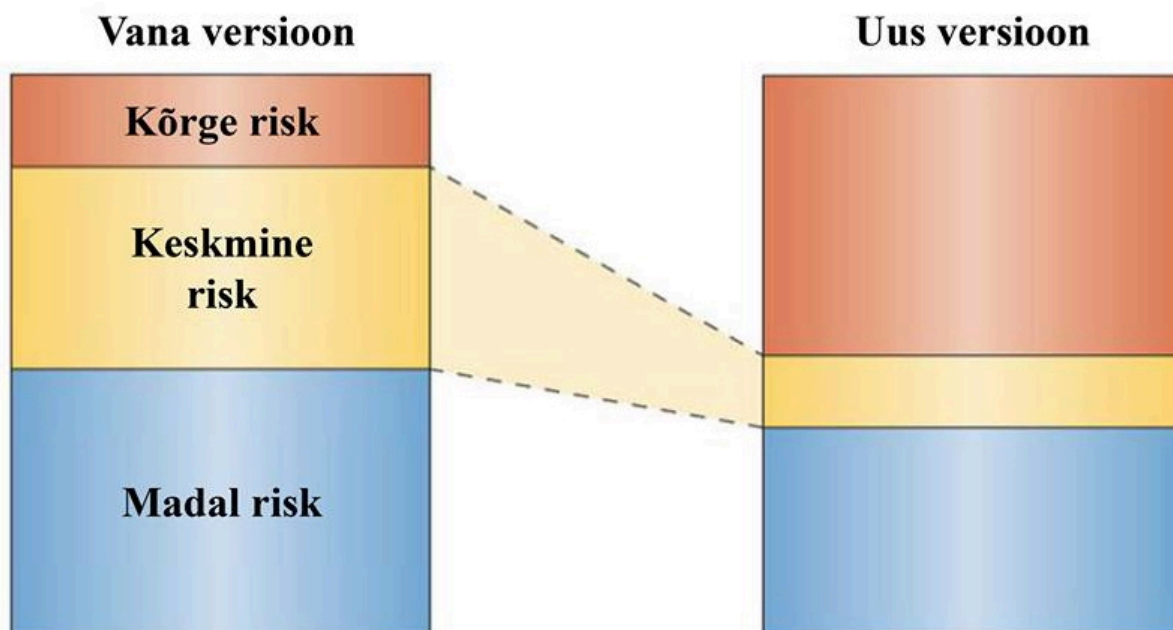
<sup>2</sup> [http://www.epi.bris.ac.uk/CVDethrisk/CHD\\_CVD\\_form.html](http://www.epi.bris.ac.uk/CVDethrisk/CHD_CVD_form.html)

näidatud, et klassikaliste riskitegurite võime hinnata südame-veresoonkonnahaiguse riski kahaneb vanusega (Kannel ja D'Agostino, 1995; Kannel, 2002), siis võib selliste mudelite rakendamine eakamate hulgas viia väärade järeldusteni (Wilson jt., 1998; Assmann jt., 2002; Conroy jt., 2003). Näiteks on leitud, et üld- ja LDL-kolesterooli tase veres on väga tugev kardiovaskulaarse riski indikaator keskealiste hulgas, kuid eakamatel on vastav seos oluliselt nõrgem (Psaty jt., 2004).

Oluline tähelepanek on veel see, et Framinghami mudelis puuduvad mitmed potentsiaalselt olulised riskitegurid: perekondlik eelsoodumus, kehamassiindeks, metaboolne sündroom, sotsiaal-majanduslik staatus ja füüsiline aktiivsus (Dent, 2010a). Kuigi ligikaudu kolm neljandikku südame isheemiatõve riskist on seletatav traditsiooniliste riskifaktoritega, tuleks efektiivsemaks hindamiseks määrata ka veres leiduvate biomarkerite tase (Dent, 2010b). Mitmed uued biomarkerid, sealhulgas C-reaktiivne valk (Folsom jt., 2006), homotsüsteiin (Ebbing jt., 2008), foolhape (Albert jt., 2008), troponiin I (Zethelius jt., 2006), troponiin T (Otsuka jt., 2010), apolipoproteiinid (Clarke jt., 2007) ning B-tüüpi natriureetiline peptiid (Di Angelantonio jt., 2009) on efektiivsed indikaatorid kõrge kardiovaskulaarse riski tuvastamiseks ning nende lisamine olemasolevatesse mudelitesse aitaks riski hindamist täpsustada. Töö nende markerite kaasamisel kliinilisse praktikasse on praegu veel poolik, kuid põhjalikult uurimise all.

Hiljuti loodi Framinghami mudelist edasiarendus, mis lubab ennustada ka ajuinsuldi tekke riski ning sisaldab uuendatud võrrandeid naiste ja afroameeriklaste kardiovaskulaarse riski hindamiseks. Kuigi modifitseeritud mudel põhineb samade riskifaktorite hindamisel, on keskmine riskirühm võrreldes eelneva versiooniga oluliselt kitsam (joonis 3). Uuendatud riskiskoori rakendamisel klassifitseeritakse peaaegu iga >60 aasta vanune mees kõrgesse riskirühma (Blaha ja Blumenthal, 2014). Sellega kaasneb suurenenud statiinide (lipiidide taset alandavad ravimid) väljakirjutamine ja tarbimine. Mitmete spetsialistide hulgas põhjustab see aga vastakaid arvamusi, sest statiinide manustamist on seostatud diabeedi kujunemisega (Carter jt., 2013).





**Joonis 3.** Uuendatud mudeli riskirühmade tulpade laius võrreldes eelneva Framinghami mudeliga (Blaha ja Blumenthal, 2014, kohandatud).

#### 1.4.2 Muud riskiskoorid

Euroopas välja töötatud riskimudel SCORE (*Systematic Coronary Risk Evaluation*) keskendub ainult fataalsetele kardiovaskulaarsetele juhtudele (näiteks müokardiinfarkt) (Grover ja Lowensteyn, 2011). Nimetatud riskiskoori arvutamisel võetakse arvesse vanus, sugu, süstoolne vererõhk, suitsetamisharjumused ning üldkolesterooli tase või üld- ja HDL-kolesterooli suhe (Conroy jt., 2003). SCORE mudelit on palju kritiseeritud, kuna see jätab arvestamata mitmed olulised riskitegurid, sealhulgas teist tüüpi diabeedi ning ülekaalulisuse (Jürgensen, 2006). SCORE mudeli eelis võrreldes Framinghami riskiskooriga on see, et SCORE töötati välja suurema ja mitmekesisema kohordi kaasamisel (Dent, 2010a).

Lisaks Framinghami ja SCORE mudelile on kasutusel veel mitmeid teisi riskiskoori. Šotimaal välja töötatud ASSIGN oli esimene, mille riski arvutamise algoritmi kaasati ka sotsiaal-majanduslik staatus ja perekondlik eelsoodumus (Woodward jt., 2007). Saksamaal arendatud PROCAM arvutab kardiovaskulaarset riski järgmiste tegurite põhjal: vanus, LDL- ja HDL-kolesterool, süstoolne vererõhk, perekondlik eelsoodumus müokardiinfarkti tekkeks, suhkruhaiguse ning triglütseriidide tase (Assmann, 2005). Reynoldsi riskiskoor, mis töötati välja eraldi naiste ja meeste jaoks, võtab lisaks traditsioonilistele riskifaktoritele arvesse ka perekondliku eelsoodumuse ning C-reaktiivse valguga taseme (Ridker jt., 2007; Ridker jt., 2008). Mitmete riskifaktoritega arvestavat QRISK-i peetakse Inglismaal aga täpsemaks kui



Framinghami (Collins ja Altman, 2009).

### 1.4.3 Geneetiline riskiskoor

Praegu kasutusel olevad riski arvutamise mudelid, mida rakendatakse kliinilises praktikas sageli sobivaima ravimeetodi valimiseks, põhinevad tavapärastel riskifaktoritel. Ometi umbes 20% indiviididest, kellel kujuneb tulevikus südame-veresoonkonnahaigus, ei oma ühtegi traditsioonilist riskitegurit, mistõttu praeguste mudelite järgi kategoriseeritakse need indiviidid madalasse riskirühma (Khot jt., 2003). Kuna kardiovaskulaarhaiguste üheks esimeseks ilminguks võib olla südame äkksurm, on mudelite täiustamine äärmiselt oluline. Seetõttu on hakatud lisaks veres ringlevatele markeritele suurt tähelepanu pöörama ka geneetilistele variatsioonidele, mille lisamine kasutusel olevatesse mudelitesse aitaks kategoriseerimist täpsustada (Thanassoulis ja Vasan, 2010).

Geneetilise riskiskoori loomiseks võib kasutada alleelide loendamisel põhinevat meetodit (*counted method*). Publitseeritud GWAS-i artiklite põhjal määratakse iga SNP-i riskialleel. Igal bialleelsel polümorfismil võib olla 0, 1 või 2 riskialleeli koopiat. Selle meetodi puhul eeldatakse, et iga SNP-l on võrdväärne efekt CAD-i kujunemisele ning geneetilise riskiskoori arvutamiseks liidetakse kõikide uuritavate SNP-de riskialleelide arv (Meigs jt., 2008; Brautbar jt., 2012). Alleelide loendamise meetodil põhineva riskiskoori arvutamiseks kasutatakse järgmist valemit:

$$\text{Loendatud riskiskoor} = \text{riskialleel}_{\text{SNP1}} + \text{riskialleel}_{\text{SNP2}} + \dots + \text{riskialleel}_{\text{SNP}_n}.$$

Alternatiivne võimalus on kaalutud riskiskoori meetod (*weighted method*), kus iga lookuse riskialleelide arv korrutatakse konkreetse SNP-i efekti suurusega ning seejärel tulemused summeeritakse (Cornelis jt., 2009). Nimetatud meetodi alusel loodud riskiskoori arvutamiseks kasutatakse järgmist valemit:

$$\text{Kaalutud riskiskoor} = \text{SNP}_1 \times \text{beta}_1 + \text{SNP}_2 \times \text{beta}_2 + \dots + \text{SNP}_n \times \text{beta}_n.$$

Davies ja kolleegide uuringu tulemusena selgus, et kaalutud geneetiline riskiskoor aitab paremini ennustada CAD-i teket kui alleelide loendamise meetodil põhinev mudel (Davies jt., 2010).

Lluís-Ganella ja kolleegide uuringus näidati, et riskialleelide arvu ja CAD-i tekke vahel on positiivselt korrelatiivne seos. See tähendab, et rohkem riskialleele kandvatel indiviididel on suurem tõenäosus südame isheemiatõve kujunemiseks (Lluís-Ganella jt., 2010). Boltoni ja

kolleegide uuringus selgus, et kaalutud geneetilise riskiskoori lisamine traditsioonilistel riskifaktoritel põhinevasse mudelisse parandas oluliselt CAD-i riski hindamist (Bolton jt., 2013). Kuigi mitmed uuringud on kinnitanud lookuse 9p21.3 ja südame isheemiatõve vahelist seost, ei aidanud nimetatud geneetilise markeri lisamine tavapäraste riskitegurite põhjal loodud mudelisse haigusrisiki paremini ennustada (Paynter jt., 2010). Samas kaalutud geneetiline riskiskoor, mis hindas lookuse 9p21.3 ja veel 11 geneetilise variandi mõju CAD-i kujunemisele, parandas oluliselt riski hindamist (Davies jt., 2010). Ripatti ja tema kolleegide poolt läbi viidud uuringus modelleeriti 13-l SNP-l põhinev kaalutud geneetiline riskiskoor koos teiste tavapäraste riskiteguritega. Nimetatud mudel parandas nende indiviidide ümberklassifitseerimist, kes traditsiooniliste riskifaktorite hindamisel olid paigutatud keskmisesse riskikategooriasse (Ripatti jt., 2010). Ka mitmetes teistes riskiskoori uuringutes on demonstreeritud indiviidide ümberklassifitseerimise paranemist geneetiliste markerite kaasamisel (Kathiresan jt., 2008; Brautbar jt., 2009).

## 2. EKSPERIMENTAALOSA

### 2.1 Töö eesmärk

Käesoleva töö praktilise osa eesmärgiks oli hinnata, kas geneetilistel markeritel põhinev kaalutud riskiskoor võimaldab ennustada südame isheemiatõve olemasolu Eesti populatsioonis.

### 2.2 Materjalid ja metoodika

#### 2.2.1 Valim

Uurimus põhines Tartu Ülikooli Eesti Geenivaramu geenidonoritel, kes on biopangaga liitunud üle kogu Eesti. Lähtudes rahvusvahelisest haiguste klassifikatsioonist (RHK-10)<sup>3</sup>, kaasati uurimusse juhud, kellel oli mõni järgmistest fenotüüpidest: stenokardia ehk rinnaangiin ehk rinnaahistus, äge müokardiinfarkt, korduv müokardiinfarkt või krooniline südame isheemiatõbi. Geneetiliseks analüüsiks kasutati nende geenidonorite DNA andmeid, mis olid kommertsiaalse Illumina HumanOmniExpress kiibiga genotüpiseeritud ning 1000 Genoomi Projekti raames sekveneeritud referentshaplotüüpide alusel imputeeritud. Valimi suuruseks oli 7599 indiviidi (4101 naist ja 3498 meest; vanuses 18 kuni 103 eluaastat). Haigusjuhte oli kokku 878 indiviidi, kontrollgrupp koosnes 6038 indiviidist ning diagnoos oli teadmata 683 inimese puhul. Valimi kirjeldus on toodud tabelis 1. Käesoleva töö raames teostatavate uuringute läbiviimiseks on geenidonorid andnud informeeritud nõusoleku ning olemas on Tartu Ülikooli inimuuringute eetika komitee kooskõlastus.

**Tabel 1.** Käesoleva töö praktilise osa läbiviimisel kasutatud valimi kirjeldus.

	Mehed	Naised
<b>Sugu</b>	46%	54%
<b>Vanus</b>	50,8 (19,5)	52,9 (21,0)
<b>CAD</b>	14%	11%

Tabelis on toodud meeste ja naiste osakaal, meeste ja naiste keskmine vanus ning sulgudes standardhälve, CAD-i esinemise osakaal valimisse kuuluvatel meestel ja naistel.

<sup>3</sup> <http://rhk.sm.ee/>

## 2.2.2 Geneetiliste markerite valik

Riskiskoori arvutamiseks valiti geneetilised markerid CAD-i fenotüübi põhistes GWAS-i publikatsioonides identifitseeritud statistiliselt oluliste ( $p$ -väärtus  $< 5 \times 10^{-8}$ ) SNP-de hulgast. Markerite valikul lähtuti kahest kriteeriumist: GWAS-i analüüs pidi olema teostatud Euroopa või Põhja-Ameerika populatsioonidel ning statistiliselt oluline leid pidi olema vähemalt ühel sõltumatul valimil replitseeritud. Sellest lähtuvalt osutusid valituks 33 geneetilist markerit, millest 13 olid meie andmestikus kiibiga genotüpiseeritud ja 20 imputeeritud. Andmed SNP-de kohta võeti GWAS-i kataloogist<sup>4</sup> (tabel 2). Tabelis toodud šansside suhteid kasutatakse riskiskoori valemis SNP-de efekti suurusena (beta), võttes neist eelnevalt naturaallogaritm.

Markerite imputeerimisel on oluline imputatsioonikvaliteet (Marchini, 2010). Tavapäraselt loetakse väärtusega  $>0,8$  imputeeritud markereid piisavalt usaldusväärseteks (Southam jt., 2011). Käesoleva töö praktilises osas kasutatud markerite imputatsioonikvaliteedi väärtus oli  $>0,9$ .

**Tabel 2.** Riskiskooride arvutamisel kasutatud 33 markerit.

SNP-i ID	Lookus	Geen või lähim geen	RAF (riskialleel)	OR (95% CI)	p-väärtus	Viide
rs1333049	9p21.3	<i>CDKN2B-AS1</i>	0,47 (C)	1,23	$1,39 \times 10^{-52}$	1
rs602633	1p13.3	<i>PSRC1</i>	0,77 (C)	1,12	$1,47 \times 10^{-25}$	1
rs17114036	1p32.2	<i>PPAP2B</i>	0,91 (A)	1,17 (1,13 – 1,22)	$4,00 \times 10^{-19}$	2
rs11556924	7q32.2	<i>ZC3HC1</i>	0,62 (C)	1,09 (1,07 – 1,12)	$9,00 \times 10^{-18}$	2
rs964184	11q23.3	<i>ZNF259</i>	0,13 (G)	1,13 (1,10 – 1,16)	$1,00 \times 10^{-17}$	2
rs579459	9q34.2	<i>ABO</i>	0,21 (C)	1,10 (1,07 – 1,13)	$4,00 \times 10^{-14}$	2
rs1412444	10q23.31	<i>LIPA</i>	0,42 (T)	1,09 (1,07 – 1,12)	$3,00 \times 10^{-13}$	3
rs7173743	15q25.1	<i>MORF4L1</i>	0,58 (T)	1,07	$6,74 \times 10^{-13}$	1
rs9818870	3q22.3	<i>MRAS</i>	0,15 (T)	1,15 (1,11 – 1,19)	$7,44 \times 10^{-13}$	4
rs12190287	6q23.2	<i>TCF21</i>	0,62 (C)	1,08 (1,06 – 1,10)	$1,00 \times 10^{-12}$	2
rs9515203	13q34	<i>COL4A2</i>	0,74 (T)	1,08	$5,85 \times 10^{-12}$	3
rs3798220	6q25.3	<i>LPA</i>	0,02 (C)	1,51 (1,33 – 1,70)	$3,00 \times 10^{-11}$	2

<sup>4</sup> <http://www.genome.gov/gwastudies/>

rs3184504	12q24.12	<i>SH2B3</i>	0,40 (T)	1,07 (1,04 – 1,10)	5,44 x 10 <sup>-11</sup>	1
rs2895811	14q32.2	<i>HHIPL1</i>	0,43 (C)	1,07 (1,05 – 1,10)	1,00 x 10 <sup>-10</sup>	2
rs1746048	10q11.21	<i>CXCL12</i>	0,87 (C)	1,09 (1,07 – 1,13)	3,00 x 10 <sup>-10</sup>	2
rs9982601	21q22.11	<i>MRPS6</i>	0,15 (T)	1,18 (1,12 – 1,24)	4,00 x 10 <sup>-10</sup>	2
rs12936587	17p11.2	<i>PEMT</i>	0,56 (G)	1,07 (1,05 – 1,09)	4,00 x 10 <sup>-10</sup>	2
rs7025486	9q33	<i>DAB2IP</i>	0,14 (A)	1,09 (1,04 – 1,15)	4,60 x 10 <sup>-10</sup>	5
rs1122608	19p13.2	<i>SMARCA4</i>	0,77 (G)	1,14 (1,09 – 1,18)	1,00 x 10 <sup>-9</sup>	2
rs12413409	10q24.32	<i>CNNM2</i>	0,89 (G)	1,12 (1,08 – 1,16)	1,00 x 10 <sup>-9</sup>	2
rs12526453	6p24.1	<i>PHACTR1</i>	0,67 (C)	1,10 (1,06 – 1,13)	1,00 x 10 <sup>-9</sup>	2
rs216172	17p13.3	<i>SMG6</i>	0,37 (C)	1,07 (1,05 – 1,09)	1,00 x 10 <sup>-9</sup>	2
rs6725887	2q33.2	<i>WDR12</i>	0,15 (C)	1,14 (1,09 – 1,19)	1,00 x 10 <sup>-9</sup>	2
rs3869109	6p21.33	<i>HCG27</i>	0,55 (G)	1,14	1,12 x 10 <sup>-9</sup>	6
rs974819	11q22.3	<i>PDGFD</i>	0,32 (T)	1,07 (1,04 – 1,09)	2,00 x 10 <sup>-9</sup>	3
rs4773144	13q34	<i>COL4A2</i>	0,44 (G)	1,07 (1,05 – 1,09)	4,00 x 10 <sup>-9</sup>	2
rs17465637	1q41	<i>MIA3</i>	0,74 (C)	1,14 (1,09 – 1,20)	1,00 x 10 <sup>-8</sup>	2
rs17609940	6p21.31	<i>ANKS1A</i>	0,75 (G)	1,07 (1,05 – 1,10)	1,00 x 10 <sup>-8</sup>	2
rs46522	17q21.32	<i>UBE2Z</i>	0,53 (T)	1,06 (1,04 – 1,08)	2,00 x 10 <sup>-8</sup>	2
rs11206510	1p32.3	<i>PCSK9</i>	0,81 (T)	1,15 (1,10 – 1,21)	2,02 x 10 <sup>-8</sup>	7
rs10953541	7q22.3	<i>BCAP29</i>	0,80 (C)	1,08 (1,05 – 1,11)	3,00 x 10 <sup>-8</sup>	3
rs2075650	19q13.32	<i>TOMM40</i>	0,14 (G)	1,14 (1,09 – 1,19)	3,20 x 10 <sup>-8</sup>	8
rs2505083	10p11.23	<i>KIAA1462</i>	0,38 (C)	1,07 (1,04 – 1,09)	4,00 x 10 <sup>-8</sup>	3

RAF – riskialleeli sagedus (*risk allele frequency*); OR – šansside suhe (*odds ratio*); CI – usaldusintervall (*confidence interval*). Viited: 1 – CARDIoGRAMplusC4D Consortium jt., 2013; 2 – Schunkert jt., 2011; 3 – Coronary Artery Disease (C4D) Genetics Consortium, 2011; 4 – Erdmann jt., 2009; 5 – Gretarsdottir jt., 2010; 6 – Davies jt., 2012; 7 – Myocardial Infarction Genetics Consortium jt., 2009; 8 – IBC 50K CAD Consortium, 2011.

### 2.2.3 Andmete analüüsitavale kujule viimine

Genotüpiseeritud ja imputeeritud markerite algandmestikust võeti analüüsimiseks välja ainult valitud 33 markeri genotüüpide info. Vastav genotüübiandmestik oli sellisel kujul, kus ridades olid SNP-de tähised ning veergudes iga indiviidi genotüüp vastavalt homo- või

heterosügootsele olekule. Kasutades Krista Fischeri poolt koostatud R-i skripti, viidi andmestik sellisele kujule, kus ridadeks olid indiviidid, veergudeks SNP-de tähised ning väärtusteks riskialleelide arv.

#### 2.2.4 Statistiline analüüs ja riskiskoori arvutamine

Andmeanalüüs teostati statistikapaketis R<sup>5</sup>, kasutades versiooni 2.15.3. Statistilise analüüsi läbiviimiseks ja riskiskoori arvutamiseks kasutati Kristi Lälli juhendamisel koostatud R-skripte. Enne riskiskooride arvutamist kontrolliti, kummalt DNA ahelalt on SNP-i riskialleel genotüübiandmestikus raporteeritud. Ahelate kontrolli puhul eeldatakse, et GWAS-i artiklites on riskialleel raporteeritud alati “+” ahelalt, kui ei ole öeldud teisiti. Seega on analüüsi kaasatavate markerite riskialleeli ja efekti suuna täpsustamiseks vaja kontrollida, et genotüübiandmestiku markerite info oleks samal kujul. Selleks võrreldi GWAS-ides raporteeritud riskialleele genotüübiandmestikuga. Kui genotüübiandmestikus oli mõni SNP tähistatud “-“ ahela nukleotiididega, siis asendati see “+” ahela omaga (näiteks kui GWAS-is on raporteeritud SNP<sub>1</sub> efektialleeliga A ja referentsalleeliga C, kuid genotüübiandmestikus on alleel T ja referentsalleeliks G, siis on selge, et kodeerimiseks on kasutatud erinevaid ahelaid). Kui genotüübiandmestikus oli riskialleeli asemel kodeeritud referentsalleel, kasutati riskiskoori arvutamiseks GWAS-is raporteeritud efektisuuna vastupidise märgiga väärtust.

Riskiskoori arvutamiseks kasutati järgmist valemit:

$$Riskiskoor = SNP_1 \times beta_1 + SNP_2 \times beta_2 + \dots + SNP_n \times beta_n,$$

kus SNP tähistab riskialleelide arvu ja beta logistilist regressioonikordajat.

Riskiskoori efekti ja olulisuse hindamisel kasutati logistilise regressiooni mudelit CAD-i jaoks, kus seletavateks tunnusteks olid vanus, sugu ja geneetiline riskiskoor.

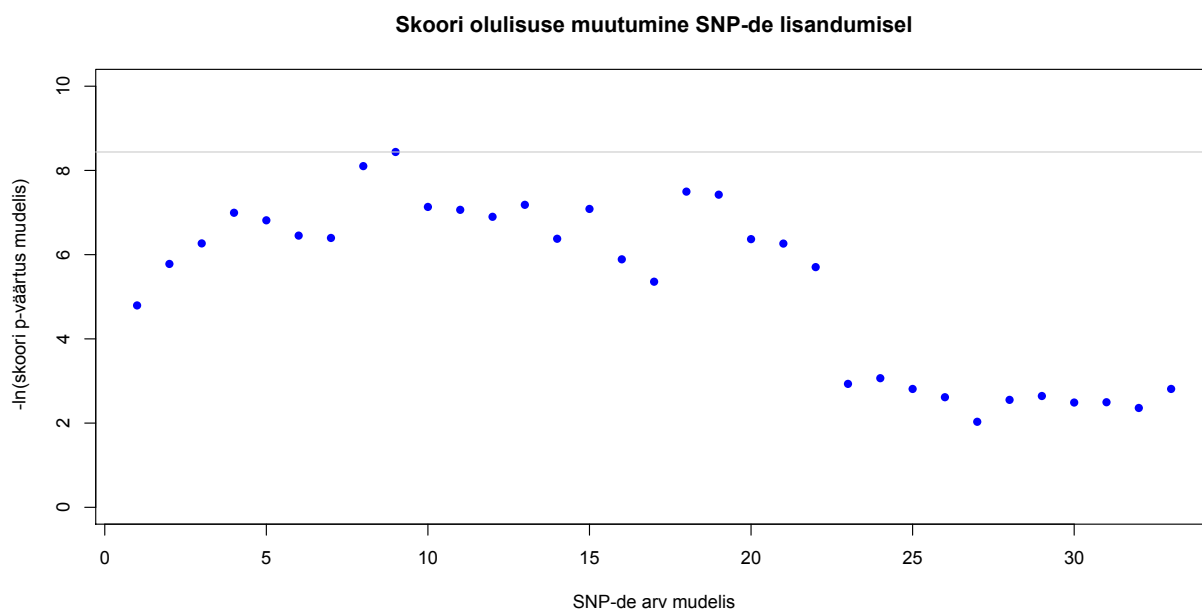
### 2.3 Tulemused

33 SNP-i põhjal arvutatud riskiskoori p-väärtuseks tuli 0,0601. Kuna ületati olulisuse nivood 0,05, ei saa väita, et nendel markeritel põhinev geneetilise riskiskoori mudel aitaks ennustada

---

<sup>5</sup> [www.R-project.org](http://www.R-project.org)

CAD-i olemasolu Eesti populatsioonis. Seetõttu otsustati uurida, kuidas SNP-d üksikuna skoori mõjutavad. Selleks lisati suureneva p-väärtuse järgi SNP-e ükshaaval mudelisse ja selgus, et parim skoor saavutatakse esimese üheksa markeriga, mille p-väärtused jäävad vahemikku  $1,39 \times 10^{-52}$  kuni  $7,44 \times 10^{-13}$ . Vastava skoori p-väärtuseks tuli 0,000216. Edasiste markerite lisamine enam skoorile midagi juurde ei andnud (joonis 4).



**Joonis 4.** Skoori olulisuse muutumine SNP-de lisamisel p-väärtuse järgi. X-teljel on toodud SNP-de arv mudelis ning y-teljel naturaallogaritm skoori p-väärtusest mudelis. Horisontaalne joon märgib kõige väiksema p-väärtusega skoori.

Kommertsiaalse ülegenoomse kiibiga genotüpiseerimisel jäävad paljud SNP-d identifitseerimata, misõttu kasutatakse analüüsitavate markerite arvu suurendamiseks imputeerimismeetodit. Kuna imputeerimine põhineb genotüpiseerimata jäänud SNP-de ennustamisel referentspaneeli alusel, ei pruugi see täpne olla (Li jt., 2009). Seetõttu otsustati uurida, kas imputeeritud markerite kasutamine mõjutab käesoleva töö uuringu tulemust. Selleks valiti SNAP programmi<sup>6</sup> ja dbSNP andmebaasi<sup>7</sup> alusel imputeeritud markerite asenduseks samast haplotüübiblokist täielikult aheldunud ( $LD=1$ ) markerid, mis on ülegenoomse kiibi peal esindatud. Seejärel teostati kaks erinevat analüüsi, et võrrelda genotüpiseeritud ja imputeeritud SNP-del arvutatud riskiskoori ainult genotüpiseeritud SNP-del põhineva skooriga. Kuna seitsmele SNP-le ei leitud asendusmarkerit, arvutati mõlemad skoorid 26 markeri põhjal (tabel 3 lisas). Imputeeritud ja genotüpiseeritud markerite põhjal

<sup>6</sup> <https://www.broadinstitute.org/mpg/snap/ldsearch.php>

<sup>7</sup> <http://www.ncbi.nlm.nih.gov/SNP/>

arvutatud skoori p-väärtuseks tuli 0,116 ning ainult genotüpiseeritud markeritel põhineva skoori p-väärtuseks tuli 0,118. Kuna mõlemal juhul ületati olulisuse nivood 0,05 ning p-väärtused olid üksteisele väga sarnased, võib järeldada, et imputeeritud markerite kasutamine analüüsitulemust ei mõjutanud.

Oluline on tähelepanek, et markeril rs3869109 on meie andmestikus tugev vastupidine efekt võrreldes publitseeritud GWAS-i tulemusega (Davies jt., 2012). Vastav marker paikneb inimese leukotsüüdi antigeeni (*human leukocyte antigen* – HLA) lookuses paiknevate geenide *HCG27* ja *HLA-C* vahel. See polümorfsete geenide rikas piirkond omab olulist rolli põletiku kujunemisel, immuunvastuse esile kutsumisel ja raku pinnal olevate koosobivusantigeenide äratundmisel. Kuna HLA lookusele toimib tugev evolutsiooniline surve, siis võib nimetatud markeri vastupidine efekt tuleneda populatsiooni etnilise tausta tõttu konkreetselt meie valimist.

## 2.4 Arutelu

Käesoleva töö praktilise osa tulemusena selgus, et kõigi uurimusse kaasatud 33 SNP-i põhjal loodud kaalutud geneetiline riskiskoor meie valimi puhul väga efektiivselt ei tööta, kuigi tulemus on piiripealne. Kõige paremini töötas üheksa väikseima p-väärtusega markeri põhjal konstrueeritud skoor. Et selgitada nende üheksa geneetilise markeri mõju südame isheemiatõve tekkes, tuleb analüüsida nende rolli bioloogilises kontekstis.

Mitmeid käesolevas töös kasutatud SNP-e on seostatud juba teadaolevate südame isheemiatõve riskifaktoritega, näiteks lipiidide metabolismiga. Mõnede markerite puhul on viidatud aga võimalikule assotsiatsioonile haiguse progressioonis või põletikulises vastuses. Kuna osa SNP-e ei ole seotud ühegi klassikalise CAD-i riskiteguriga, tuleb nende markerite täpse toimimise mehhanismi ja rolli väljaselgitamiseks haiguse kujunemisel lookusi põhjalikumalt uurida.

Lookuses 1p13.3 paiknev SNP rs602633 asub geenide *PSRC1* ja *CELSR2* vahel, mis kodeerivad vastavalt mikrotuubulite destabiliseerimises osalevat valku (Hsieh jt., 2007) ning kontakt-vahendatud raku adhesioonis ja retseptor-ligand interaktsioonides funktsioneerivat proteiini (Vincent jt., 2000). Lisaks sellele markerile on nimetatud lookuses näidatud veel SNP-ga rs602633 täielikult aheldunud ( $LD=1$ ) geneetilise markeri rs599839 seost suurenenud



CAD-i riski ning üld- ja LDL-kolesterooolitasemega (Samani jt., 2008). Kuigi markeri rs599839 põhjuslikku efekti haiguse kujunemisele ning kolesterooli tasemele pole veel selgitatud, viitab mitme markeri assotsiatsioon ühes lookuses asjaolule, et GWAS-is identifitseeritud väikseima p-väärtusega markerid ei pruugi otseselt olla haigust põhjustavad variandid. Põhjuslik SNP võib olla tuvastatud markeritega aheldunud ning omada palju kõrgemat p-väärtust (Stranger jt., 2011). Teine lipiidide metabolismiga seotud SNP rs964184 asub lookuses 11q23.3. Nimetatud markeri lähim geen *ZNF259* kodeerib valku ZPR1, mis on oluline reguleeriv proteiin normaalseks tuumakese talitlemiseks raku proliferatsioonis ja signaaliülekanDES (Galcheva-Gargova jt., 1998). On näidatud, et see valk suurendab triglütseriidide taset ja lipoproteiiniga seostunud fosfolipasaas A2 (*lipoprotein-associated phospholipase A2* – Lp-PLA2) aktiivsust (Willer jt., 2008). Lp-PLA2, mida peetakse sõltumatuks kardiovaskulaarseks riskiteguriks ja põletikuliseks biomarkeriks (Rubinstein ja Izhakov, 2011), omab olulist rolli ateroskleroosi arengus (Zheng jt., 2011), mõjutades LDL-partikleid (Suchindran jt., 2010). Lisaks on erineva etnilise taustaga populatsioonides leitud seos geenis *LIPA* asuva SNP-i rs1412444 ja südame isheemiatõve vahel (Wild jt., 2011; Vargas-Alarcón, 2013). Nimetatud geneetilist markerit on seostatud suurenenud *LIPA* ekspressioonitasemega (Wild jt., 2011). Geen *LIPA* kodeerib LAL (*lysosomal acid lipase*) valku, mis hüdrolüüsib raku lüsoosoomides kolesterooliestreid ja triglütseriide vaba kolesterooli ja rasvhapete saamiseks. Kui vastav ensüüm normaalselt ei funktsioneer, võivad lipiidid rakus akumulēeruda ja põhjustada vahtrakkude ning seejärel aterosklerootilise naastu moodustumist (Zschenker jt., 2006). Seega viitavad nimetatud lookused lipiidide metabolismi osalusele kui väga tähtsale protsessile ateroskleroosi patogeneesis.

Geeni *ABO* lähedal asuv SNP rs579459 on seotud LDL-kolesterooli (Chasman jt., 2009), ICAM-1 (Paré jt., 2008) ning E- (Paterson jt., 2009) ja P-selektiinide (Barbalic jt., 2010) tasemega. ICAM-1 ning E- ja P-selektiinid osalevad põletikulistes protsessides, edendades leukotsüütide adhesiooni veresoone seinale endoteelirakkudele (Barbalic jt., 2010). Seega võib oletada, et lookuses 9q34.2 olevatel variantidel on oluline funktsioon ateroskleroosi põletikulise vastuse kujunemises. Teine inflammasiooniga seotud SNP rs9818870, mis aitab ennustada kardiovaskulaarse haiguse riski tervetel indiviididel, paikneb lookuses 3q22.3 (Ellis jt., 2011). Nimetatud marker asub geeni *MRAS* 3'UTR (3'*untranslated region*) regioonis mikro-RNA seondumiskohtade lähedal. 33 kiloaluspaari pikkuselt *MRAS* geenilt kodeeritakse GTP-d siduvate Ras valkude superperekonda kuuluvat proteiini M-Ras, mida ekspresseeritakse suures koguses südames (Erdmann jt., 2009). Hiirtega läbiviidud uuringutes

on demonstreeritud M-Ras valgu osalemist immuunrakkude adhesiooni signalisatsioonis (Yoshikawa jt., 2007). Oluline oleks uurida, kuidas SNP nimetatud geenile efekti avaldab ning immuunrakkude adhesiooni signalisatsiooni mõjutab. Lookuse detailsem uurimine võimaldaks ateroskleroosi põletikuliste protsesside kujunemist paremini selgitada.

Geenis *ZC3HCl* asuvat geneetilist markerit rs11556924 on seostatud unearteri intima-media paksenemisega reumatoidartriiti põdevatel patsientidel (López-Mejías jt., 2013). Reumatoidartriit on kompleksne põletikuline haigus, mis on seotud suurenenud kardiovaskulaarhaiguste riskiga (Chung jt., 2005). Geen *ZC3HCl* kodeerib samanimelist valku, mille funktsiooniks on kontrollida mitootilise jagunemise ajastust ja kompleksis konstitutiivselt aktiivsete onkogeensete proteiinidega soodustada kartsinogeneesi arengut (Li ja Morris, 2008). Muutused *ZC3HCl* valgu stabiilsuses ja funktsionaalsetes omadustes võivad põhjustada endoteeli väärtalitlust (López-Mejías jt., 2013). Seega võib nimetatud markeril olla oluline roll ateroskleroosi kujunemisel, kuna häired endoteeli läbilaskevõimes soodustavad haiguse teket. Ateroskleroosi kulu halvenemisega seotud SNP rs7173743 paikneb geenivahelises piirkonnas. Lisaks geenile *MORF4L1* paikneb nimetatud markeri lähedal ka geen *ADAMTS7*. Geeniproducti akumulierumist on näidatud pärg- ja unearteri ateroskleroosilise naastu koostises olevates silelihasrakkudes (Pu jt., 2013). *ADAMTS7* vahendab ekstratsellulaarse maatriksi valgu COMP (*cartilage oligomeric matrix protein*) degradatsiooni, mis põhjustab vaskulaarsete SMC-de migratsiooni ja intima paksenemisest pärast veresoonte kahjustust (Wang jt., 2009). Nimetatud SNP-i efekti *ADAMTS7* geenile tuleks veel uurida.

Lookuses 9p21.3 paiknevatest SNP-dest on olulisimaks osutunud rs1333049, mida on erinevates populatsioonides CAD-i fenotüübiga ka edukalt replitseeritud (Mendonça jt., 2011; Shanker jt., 2014). Oluline on märkida, et nimetatud lookuses paiknevad markerid ei ole seotud ühegi tavapärase südame isheemiatõve riskifaktoriga, mis viitab seni avastamata molekulaarsetele radadele ateroskleroosi kulus. On tõenäoline, et 9p21.3 lookuses paiknevatel geneetilistel variantidel on regulatoorne roll valke kodeerivate geenide ekspressioonil (Roberts ja Stewart, 2012b). SNP-i rs1333049 lähimate valke kodeerivate geenide *CDKN2A* ja *CDKN2B* täpse rolli ja mehhanismi väljaselgitamine haiguse kujunemisel vajab täiendavaid funktsionaalseid uuringuid. Samuti on geen *PPAP2B* paikneva SNP-i rs17114036 täpse toimimise mehhanismi mõistmiseks vaja teha funktsionaalseid katseid (Erbilgin jt., 2013).

Kuna senini identifitseeritud geneetilised markerid selgitavad ainult väikese osa südame isheemiatõve pärilikkusest, aitaks uute SNP-de tuvastamine paremini mõista haiguse geneetilist tagapõhja. Käesoleva töö praktiline osa põhines ainult geneetilistel markeritel, kuid mitmetes uuringutes on demonstreeritud, et võttes lisaks SNP-dele arvesse ka traditsioonilised riskitegurid, on haiguse kujunemise tõenäosuse hindamine oluliselt efektiivsem (Davies jt., 2010; Ripatti jt., 2010; Bolton jt., 2013). Antud töös selgus, et vähemate, kuid väiksemate p-väärtustega SNP-de kasutamine riskiskoori arvutamisel annab parema tulemuse kui rohkemate ja seejuures suuremate p-väärtustega markerite kaasamine. Samas on Paynteri ja kolleegide poolt läbiviidud uuringus selgunud, et ainult statistiliselt kõige olulisemat markerit sisaldava lookuse lisamine mudelisse ei paranda CAD-i riski hindamist (Paynter jt., 2010). Kuna südame isheemiatõve näol on tegemist äärmiselt kompleksse fenotüübiga, on uuringusse kaasatavate geneetiliste markerite valimine väga keeruline, kuid väärib tulevikus kindlasti edasist uurimist.

Enamikus riskiskooridel põhinevates töödes, kus on uuritud geneetiliste markerite ja tavapäraste riskitegurite võimet hinnata kardiovaskulaarse haigusjuhu tekke riski, on kõik SNP-d otse genotüpiseeritud. Käesolevas töös kasutati selle asemel ülegenoomse kiibi andmeid. Kuigi antud töös analüüsiti imputeerimise efekti, ei olnud seda võimalik teha seitsme SNP-i puhul, millele ei leitud asendusmarkerit tingimusel  $LD=1$ . Seetõttu tuleks need SNP-d üle genotüpiseerida.

Täpsema riski hindamise saavutamiseks tuleks käesolevas töös kasutatud geneetilised markerid lisada olemasolevatesse tavapärastel riskiteguritel põhinevatesse mudelitesse.

## KOKKUVÕTE

Südame isheemiatõbi ja sellest tingitud tüsistused põhjustavad igal aastal miljonite inimeste surma. Nimetatud haiguse puhul on tegemist äärmiselt kompleksse fenotüübiga, mille kujunemisel mängivad rolli nii geneetilised kui ka keskkonnast tingitud tegurid. Seetõttu ei ole senini suudetud luua efektiivset ravimeetodit südame isheemiatõve esinemissageduse vähendamiseks. Kuigi konkreetse tõve eest vastutavate tavapäraste riskifaktorite kindlakstegemine ei ole viinud soovitud eesmärkideni, võib ülegenoomsetes assotsiatsiooniuringutes tuvastatud geneetiliste variantide uurimine aidata indiviidide haigestumisriski parema hindamisstrateegia väljatöötamisel.

Käesolevas töös on kirjeldatud põhilisi südame isheemiatõve geneetiliste aluste selgitamiseks kasutatavaid meetodeid ning haiguse ennetamiseks või varaseks avastamiseks rakendatavaid riskiskooridel põhinevaid mudeleid. Praktilise osa eesmärgiks oli uurida, kas geneetiline riskiskoor aitab ennustada südame isheemiatõve olemasolu Eesti populatsioonis. Selleks analüüsiti GWAS-ides raporteeritud statistiliselt oluliste SNP-de efekti Eesti Geenivaramu valimis, kasutades kaalutud meetodil põhinevat geneetilist riskiskoori.

Antud töös selgus, et geneetilised markerid aitavad ennustada südame isheemiatõve olemasolu Eesti Geenivaramu valimis. Parima tulemuse andsid väikseima p-väärtusega SNP-d, millel võib olla südame-veresoonkonnahaiguse kujunemisele suurem efekt.

Kuna traditsioonilistel riskiteguritel põhinevad riskiskoori mudelid ei suuda efektiivselt ennustada kardiovaskulaarse haiguse tekke tõenäosust, võib arvata, et geneetiliste markerite lisamine olemasolevatesse mudelitesse aitaks riski hindamist oluliselt parandada.

## **Risk estimation of coronary artery disease based on genetic markers in Estonian Genome Center cohort**

Anett Tähiste

### **Summary**

Coronary artery disease (CAD) is the leading cause of death and disability worldwide, and its prevalence is expected to increase in the coming years. CAD is a multifactorial disease involving both genetic and environmental factors and is primarily caused due to a process of progressive damage of large and medium-sized coronary arteries called atherosclerosis. The techniques used to study the genetic basis of cardiovascular diseases have evolved from linkage studies to candidate gene analysis and genome-wide association studies (GWAS). Currently, approximately 40 single nucleotide polymorphisms (SNPs) associated with CAD have been identified by GWAS.

The major cardiovascular risk factors, namely male sex, hypertension, cholesterol level and smoking behaviour have been well known for a long time. Based on these factors, a number of risk prediction algorithm scores have been developed, including the Framingham risk score, that provides an estimate of the 10-year risk. However, current assessment strategies do not fully explain the cardiovascular risk. For example, it is suggested that approximately 20% of patients with CAD lack any of the conventional risk factors. This implies that additional factors might help improve the accuracy of disease risk delineation. This hypothesis has led to considerable research on nontraditional risk factors, including genetic markers, which could significantly improve risk estimation.

The aim of this study was to estimate whether weighted risk score based on genetic markers could predict the presence of coronary artery disease in Estonian population. The study involved *in silico* data from the Estonian Genome Center, Tartu University on 878 CAD cases and 6038 controls. The genetic risk score calculated as a weighted sum of CAD risk alleles was based on 33 top GWAS markers.

The genetic risk score model based on 33 SNPs, sex and age did not work well in our sample (p-value 0.0601), but the model including 9 markers with smallest p-values improved the estimation of the presence of the disease significantly (p-value 0.000216). Involvement of

imputed markers did not affect the results of our study. Risk prediction should improve considerably if traditional risk factors are taken into account as well.

Because risk score models based on traditional risk factors cannot effectively predict the likelihood of developing cardiovascular disease, it can be assumed, that inclusion of genetic markers into existing risk assessment models would greatly improve risk estimation.

# TÄNUSÕNAD

Kõige suuremad tänusõnad kuuluvad minu juhendajale, Maris Alverile, kelle abi, õpetussõnad ning panus sellesse töösse on olnud hindamatu väärtusega. Parim juhendaja, kelle toetus töö valmimisel süstis minusse motivatsiooni ja kindlustunnet, et tegelen õige asjaga.

Lisaks soovin siiralt tänada oma kaasjuhendajat Kristi Lälli, kelle abi ja kannatlikkus statistiliste mõistete selgitamisel ning andmeanalüüsi läbiviimisel olid töö valmimise seisukohalt asendamatud.

Suured tänud ka teistele biotehnoloogia õppetooli inimestele, kes olid alati nõus küsimustele vastama ja probleemide esinemise korral aitama.

Soovin tänada ka oma vanemaid, vanavanemaid ja elukaaslast, kelle moraalne toetus ning usk minusse aitasid töö valmimisele oluliselt kaasa.

## KIRJANDUSE LOETELU

- Abbate, R., Sticchi, E., Fatini, C. (2008). Genetics of cardiovascular disease. *Clin Cases Miner Bone Metab.* 5(1): 63-66.
- Albert, C.M., Cook, N.R., Gaziano, J.M., et al. (2008). Effect of folic acid and B vitamins on risk of cardiovascular events and total mortality among women at high risk for cardiovascular disease: a randomized trial. *JAMA.* 299(17): 2027-2036.
- Assmann, G. (2005). Calculating global risk: the key to intervention. *Eur Heart J. Suppl* 7: F9-14.
- Assmann, G., Cullen, P., Schulte, H. (2002). Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Münster (PROCAM) study. *Circulation.* 105(3): 310-315.
- Barbalic, M., Dupuis, J., Dehghan, A., et al. (2010). Large-scale genomic studies reveal central role of ABO in sP-selectin and sICAM-1 levels. *Hum Mol Genet.* 19(9): 1863-1872.
- Baudhuin, L.M. (2009). Genetics of coronary artery disease: focus on genome-wide association studies. *Am J Transl Res.* 1(3): 221-234.
- Beswick, A., Brindle, P. (2006). Risk scoring in the assessment of cardiovascular risk. *Curr Opin Lipidol.* 17: 375-386.
- Blahe, M.J., Blumenthal, R.S. (2014). Risk factors: New risk-assessment guidelines – more or less personalized? *Nat Rev Cardiol.* 11: 136-137.
- Boffa, M.B., Marcovina, S.M., Koschinsky, M.L. (2004). Lipoprotein(a) as a risk factor for atherosclerosis and thrombosis: mechanistic insights from animal models. *Clin Biochem.* 37(5): 333-343.
- Bolton, J.L., Stewart, M.C.W., Wilson, J.F., et al. (2013). Improvement in prediction of coronary heart disease risk over conventional risk factors using SNPs identified in genome-wide association studies. *PLoS One.* 8(2): e57310.
- Brautbar, A., Ballantyne, C.M., Lawson, K., et al. (2009). Impact of adding a single allele in the 9p21 locus to traditional risk factors on reclassification of coronary heart disease risk and implications for lipid-modifying therapy in the Atherosclerosis Risk in Communities study. *Circ Cardiovasc Genet.* 2(3): 279-285.
- Brautbar, A., Pompeii, L.A., Dehghan, A., et al. (2012). A genetic risk score based on direct associations with coronary heart disease improves coronary heart disease risk prediction in the Atherosclerosis Risk in Communities (ARIC), but not in the Rotterdam and Framingham Offspring, Studies. *Atherosclerosis.* 223(2): 421-426.
- Bronte-Stewart, B., Botha, M.C., Krut, L.H. (1962). ABO blood groups in relation to ischaemic heart disease. *Br Med J.* 1(5293): 1646-1650.
- Businaro, R., Tagliani, A., Buttari, B., et al. (2012). Cellular and molecular players in the atherosclerotic plaque progression. *Ann NY Acad Sci.* 1262: 134-141.
- CARDIoGRAMplusC4D Consortium, Deloukas, P., Kanoni, S., et al. (2013). Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet.* 45(1): 25-33.
- Carter, A.A., Gomes, T., Camacho, X., et al. (2013). Risk of incident diabetes among patients treated with statins: population based study. *BMJ.* 346: f2610.
- Cassar, A., Holmes, D.R.Jr., Rihal, C.S., et al. (2009). Chronic coronary artery disease: diagnosis and management. *Mayo Clin Proc.* 84(12): 1130-1146.
- Castelli, W.P. (1996). Lipids, risk factors and ischaemic heart disease. *Atherosclerosis. Suppl* 124: S1-9.
- Catapano, A.L., Maggi, F.M., Tragni, E. (2000). Low density lipoprotein oxidation, antioxidants, and atherosclerosis. *Curr Opin Cardiol.* 15(5): 355-363.



- Celermajer, D.S., Chow, C.K., Marijon, E., et al. (2012). Cardiovascular disease in the developing world: prevalences, patterns, and the potential of early disease detection. *J Am Coll Cardiol.* 60(14): 1207-1216.
- Chasman, D.I., Pare, G., Mora, S., et al. (2009). Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. *PLoS Genet.* 5(11): e1000730.
- Chen, S.N., Ballantyne, C.M., Gotto, A.M.Jr., et al. (2005). A common PCSK9 haplotype, encompassing the E670G coding single nucleotide polymorphism, is a novel genetic marker for plasma low-density lipoprotein cholesterol levels and severity of coronary atherosclerosis. *J Am Coll Cardiol.* 45(10): 1611-1619.
- Chen, S.N., Ballantyne, C.M., Gotto, A.M.Jr., et al. (2009). The 9p21 susceptibility locus for coronary artery disease and the severity of coronary atherosclerosis. *BMC Cardiovasc Disord.* 9: 3.
- Chhabra, N. (2009). Endothelial dysfunction – a predictor of atherosclerosis. *IJMU.* 4(1): 33-41.
- Chial, H. (2008). Mendelian genetics: patterns of inheritance and single-gene disorders. *Nat Education.* 1(1): 63.
- Chung, C.P., Oeser, A., Raggi, P., et al. (2005). Increased coronary-artery atherosclerosis in rheumatoid arthritis: relationship to disease duration and cardiovascular risk factors. *Arthritis Rheum.* 52(10): 3045-3053.
- Civeira, F., International Panel on Management of Familial Hypercholesterolemia. (2004). Guidelines for the diagnosis and management of heterozygous familial hypercholesterolemia. *Atherosclerosis.* 173(1): 55-68.
- Clarke, R., Emberson, J.R., Parish, S., et al. (2007). Cholesterol fractions and apolipoproteins as risk factors for heart disease mortality in older men. *Arch Intern Med.* 167: 1373-1378.
- Collins, G.S., Altman, D.G. (2009). An independent external validation and evaluation of QRISK cardiovascular risk prediction: a prospective open cohort study. *BMJ.* 339: b2584.
- Conroy, R.M., Pyörälä, K., Fitzgerald, A.P., et al. (2003). Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J.* 24(11): 987-1003.
- Cooper, R., Cutler, J., Desvigne-Nickens, P., et al. (2000). Trends and disparities in coronary heart disease, stroke, and other cardiovascular diseases in the United States: findings of the national conference on cardiovascular disease prevention. *Circulation.* 102(25): 3137-3147.
- Cornelis, M.C., Qi, L., Zhang, C., et al. (2009). Joint effects of common genetic variants on the risk for type 2 diabetes in U.S. men and women of european ancestry. *Ann Intern Med.* 150(8): 541-550.
- Coronary Artery Disease (C4D) Genetics Consortium. (2011). A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. *Nat Genet.* 43(4): 339-344.
- D'Agostino, R.B.Sr., Grundy, S., Sullivan, L.M., et al. (2001). Validation of the Framingham coronary heart disease prediction scores: Results of a multiple ethnic groups investigation. *JAMA.* 286(2): 180-187.
- Dahlöf, B. (2010). Cardiovascular disease risk factors: epidemiology and risk assessment. *Am J Cardiol.* 105(1): 3A–9A.
- Damgaard, D., Jensen, J.M., Larsen, M.L., et al. (2004). NO genetic linkage or molecular evidence for involvement of the PCSK9, ARH or CYP7A1 genes in the familial hypercholesterolemia phenotype in a sample of Danish families without pathogenic mutations in the LDL receptor and *APOB* genes. *Atherosclerosis.* 177(2): 415-422.

- Davies, R.W., Dandona, S., Stewart, A.F., et al. (2010). Improved prediction of cardiovascular disease based on a panel of single nucleotide polymorphisms identified through genome-wide association studies. *Circ Cardiovasc Genet*. 3(5): 468-474.
- Davies, R.W., Wells, G.A., Stewart, A.F., et al. (2012). A genome-wide association study for coronary artery disease identifies a novel susceptibility locus in the major histocompatibility complex. *Circ Cardiovasc Genet*. 5(2): 217-225.
- Davis, M., Andrade, J., Taylor, C., et al. (2010). Cardiovascular risk factors and models of risk prediction: Recognizing the leadership of Dr Roy Dawber. *BCM J*. 52(7): 342-348.
- Dent, T.H. (2010a). Predicting the risk of coronary heart disease I. The use of conventional risk markers. *Atherosclerosis*. 213(2): 345-351.
- Dent, T.H. (2010b). Predicting the risk of coronary heart disease. II: The role of novel molecular biomarkers and genetics in estimating risk, and the future of risk prediction. *Atherosclerosis*. 213(2): 352-362.
- Di Angelantonio E., Chowdhury, R., Sarwar, N., et al. (2009). B-type natriuretic peptides and cardiovascular risk: systematic review and meta-analysis of 40 prospective studies. *Circulation*. 120(22): 2177-2187.
- Dietter, J., Spiegel, A., an Mey, D., et al. (2004). Efficient two-trait-locus linkage analysis through program optimization and parallelization: application to hypercholesterolemia. *Eur J Human Genet*. 12(7): 542-550.
- Ding, H., Xu, Y., Wang, X., et al. (2009). 9p21 is a shared susceptibility locus strongly for coronary artery disease and weakly for ischemic stroke in Chinese Han population. *Circ Cardiovasc Genet*. 2(4): 338-346.
- Ebbing, M., Bleie, Ø., Ueland, P.M., et al. (2008). Mortality and cardiovascular events in patients treated with homocysteine-lowering B vitamins after coronary angiography: a randomized controlled trial. *JAMA*. 300(7): 795-804.
- Ellis, K.L., Frampton, C.M., Pilbrow, A.P., et al. (2011). Genomic risk variants at 1p13.3, 1q41, and 3q22.3 are associated with subsequent cardiovascular outcomes in healthy controls and in established coronary artery disease. *Circ Cardiovasc Genet*. 4(6): 636-646.
- Elosua, R., Lluís-Ganella, C., Lucas, G. (2009). Research into the genetic component of heart disease: from linkage studies to genome-wide genotyping. *Rev Esp Cardiol*. 9: 24B-38B.
- Erbilgin, A., Civelek, M., Romanoski, C.E., et al. (2013). Identification of CAD candidate genes in GWAS loci and their expression in vascular cells. *J Lipid Res*. 54(7): 1894-1905.
- Erdmann, J., Grosshennig, A., Braund, P.S., et al. (2009). New susceptibility locus for coronary artery disease on chromosome 3q22.3. *Nat Genet*. 41(3): 280-282.
- Folsom, A.R., Chambless, L.E., Ballantyne, C.M., et al. (2006). An assessment of incremental coronary risk prediction using c-reactive protein and other novel risk markers: the atherosclerosis risk in communities study. *Arch Intern Med*. 166(13): 1368-1373.
- Fox, C.S., Cupples, L.A., Chazaro, I., et al. (2004). Genomewide linkage analysis for internal carotid artery intimal medial thickness: evidence for linkage to chromosome 12. *Am J Hum Genet*. 74(2): 253-261.
- Galcheva-Gargova, Z., Gangwani, L., Konstantinov, K.N., et al. (1998). The cytoplasmic zinc finger protein ZPR1 accumulates in the nucleolus of proliferating cells. *Mol Biol Cell*. 9(10): 2963-2971.
- Ghattas, A., Griffiths, H.R., Devitt, A., et al. (2013). Monocytes in coronary artery disease and atherosclerosis: where are we now? *J Am Coll Cardiol*. 62(17): 1541-1551.
- Glass, C.K., Witztum, J.L. (2001). Atherosclerosis: the road ahead. *Cell*. 104(4): 503-516.
- Gorroochurn, P., Hodge, S.E., Heiman, G.A., et al. (2007). Non-replication of association studies: "pseudo-failures" to replicate? *Genet Med*. 9(6): 325-331.

- Greaves, D.R., Channon, K.M. (2002). Inflammation and immune responses in atherosclerosis. *Trends Immunol.* 23(11): 535-541.
- Gretarsdottir, S., Baas, A.F., Thorleifsson, G., et al. (2010). Genome-wide association study identifies a sequence variant within the *DAB2IP* gene conferring susceptibility to abdominal aortic aneurysm. *Nat Genet.* 42: 692-697.
- Grover, S.A., Lowensteyn, I. (2011). The challenges and benefits of cardiovascular risk assessment in clinical practice. *Can J Cardiol.* 27(4): 481-487.
- Hansson, G.K. (2005). Inflammation, atherosclerosis, and coronary artery disease. *N Engl J Med.* 352(16): 1685-1695.
- Helgadottir, A., Manolescu, A., Thorleifsson, G., et al. (2004). The gene encoding 5-lipoxygenase activating protein confers risk of myocardial infarction and stroke. *Nat Genet.* 36(3): 233-239.
- Helgadottir, A., Thorleifsson, G., Magnusson, K.P., et al. (2008). The same sequence variant on 9p21 associates with myocardial infarction, abdominal aortic aneurysm and intracranial aneurysm. *Nat Genet.* 40(2): 217-224.
- Helgadottir, A., Thorleifsson, G., Manolescu, A., et al. (2007). A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science.* 316(5830): 1491-1493.
- Hinohara, K., Nakajima, T., Takahashi, M., et al. (2008). Replication of the association between a chromosome 9p21 polymorphism and coronary artery disease in Japanese and Korean populations. *J Hum Genet.* 53: 357-359.
- Hozawa, A., Folsom, A.R., Sharrett, A.R., et al. (2007). Absolute and attributable risks of cardiovascular disease incidence in relation to optimal and borderline risk factors: comparison of African American with white subjects—Atherosclerosis Risk in Communities Study. *Arch Intern Med.* 167(6): 573-579.
- Hsieh, P.C., Chang, J.C., Sun, W.T., et al. (2007). p53 downstream target DDA3 is a novel microtubule-associated protein that interacts with end-binding protein EB3 and activates beta-catenin pathway. *Oncogene.* 26(34): 4928-4940.
- IBC 50K CAD Consortium. (2011). Large-scale gene-centric analysis identifies novel variants for coronary artery disease. *PLoS Genet.* 7(9): e1002260.
- International HapMap Consortium, Frazer, K.A., Ballinger, D.G., et al. (2007). A second generation human haplotype map of over 3.1 million SNPs. *Nature.* 449(7164): 851-861.
- Jürgensen, J.S. (2006). The value of risk scores. *Heart.* 92(12): 1713-1714.
- Kamstrup, P.R., Tybjaerg-Hansen, A., Steffensen, R., et al. (2009). Genetically elevated lipoprotein(a) and increased risk of myocardial infarction. *JAMA.* 301(22): 2331-2339.
- Kanjilal, S., Rao, V.S., Mukherjee, M., et al. (2008). Application of cardiovascular disease risk prediction models and the relevance of novel biomarkers to risk stratification in Asian Indians. *Vasc Health Risk Manag.* 4(1): 199-211.
- Kannel, W.B. (2002). Coronary heart disease risk factors in the elderly. *Am J Geriatr Cardiol.* 11(2): 101-107.
- Kannel, W.B., D'Agostino, R.B. (1995). The importance of cardiovascular risk factors in the elderly. *Am J Geriatr Cardiol.* 4(2): 10-23.
- Kathiresan, S., Melander, O., Anevski, D., et al. (2008). Polymorphisms associated with cholesterol and risk of cardiovascular events. *N Engl J Med.* 358(12): 1240-1249.
- Kathiresan, S., Srivastava, D. (2012). Genetics of human cardiovascular disease. *Cell.* 148(6): 1242-1257.
- Khot, U.N., Khot, M.B., Bajzer, C.T., et al. (2003). Prevalence of conventional risk factors in patients with coronary heart disease. *JAMA.* 290(7): 898-904.
- Kjolby, M., Andersen, O.M., Breiderhoff, T., et al. (2010). Sort1, encoded by the cardiovascular risk locus 1p13.3, is a regulator of hepatic lipoprotein export. *Cell*

- Metab. 12(3): 213-223.
- Kruglyak, L. (2005). Power tools for human genetics. *Nat Genet.* 37(12): 1299-1300.
- Kullo, I.J., Turner, S.T., Kardia, S.L., et al. (2006). A genome-wide linkage scan for ankle-brachial index in African American and non-Hispanic white subjects participating in the GENOA study. *Atherosclerosis.* 187(2): 433-438.
- Kumar, J., Yumnam, S., Basu, T., et al. (2011). Association of polymorphisms in 9p21 region with CAD in North Indian population: replication of SNPs identified through GWAS. *Clin Genet.* 79(6): 588-593.
- Kwon, J.M., Goate, A.M. (2000). The candidate gene approach. *Alcohol Res Health.* 24(3): 164-168.
- Lange, L.A., Lange, E.M., Bielak, L.F., et al. (2002). Autosomal genome-wide scan for coronary artery calcification loci in sibships at high risk for hypertension. *Arterioscler Thromb Vasc Biol.* 22(3): 418-423.
- Li, R., Morris, S.W. (2008). Development of anaplastic lymphoma kinase (ALK) small-molecule inhibitors for cancer therapy. *Med Res Rev.* 3: 372-412.
- Li, Y., Willer, C., Sanna, S., et al. (2009). Genotype imputation. *Annu Rev Genom Hum G.* 10: 387-406.
- Libby, P. (2012). Inflammation in atherosclerosis. *Arterioscler Thromb Vasc Biol.* 32(9): 2045-2051.
- Libby, P., Ridker, P.M., Hansson, G.K. (2011). Progress and challenges in translating the biology of atherosclerosis. *Nature.* 473(7347): 317-325.
- Libby, P., Theroux, P. (2005). Pathophysiology of coronary artery disease. *Circulation.* 111: 3481-3488.
- Lloyd-Jones, D.M. (2010). Cardiovascular risk prediction: basic concepts, current status, and future directions. *Circulation.* 121(15): 1768-1777.
- Lluís-Ganella, C., Lucas, G., Subirana, I., et al. (2010). Additive Effect of Multiple Genetic Variants on the Risk of Coronary Artery Disease. *Rev Esp Cardiol.* 63(8): 925-933.
- Lobo, I. (2008). Multifactorial inheritance and genetic disease. *Nat Education.* 1(1): 5
- López-Mejías, R., Genre, F., García-Bermúdez, M., et al. (2013). The ZC3HC1 rs11556924 polymorphism is associated with increased carotid intima-media thickness in patients with rheumatoid arthritis. *Arthritis Res Ther.* 15(5): R152.
- Lusis, A.J., Mar, R., Pajukanta, P. (2004). Genetics of atherosclerosis. *Annu Rev Genom Hum G.* 5: 189-218.
- Maas, A.H., Appelman, Y.E. (2010). Gender differences in coronary heart disease. *Neth Heart J.* 18(12): 598-602.
- Marchini, J. (2010). SNPTEST v2. Technical Details. Lk 8, valem 50.
- Marks, D., Thorogood, M., Neil, H.A., et al. (2003). A review on the diagnosis, natural history, and treatment of familial hypercholesterolaemia. *Atherosclerosis.* 168(1): 1-14.
- McCarthy, M.I., Abecasis, G.R., Cardon, L.R., et al. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet.* 9(5): 356-369.
- McGinn, T.G., Guyatt, G.H., Wyer, P.C., et al. (2000). Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. *JAMA.* 284(1): 79-84.
- McPherson R., Pertsemlidis A., Kavaslar N., et al. (2007). A common allele on chromosome 9 associated with coronary heart disease. *Science.* 316(5830): 1488-1491.
- Mecchia, D., Lavezzi, A.M., Mauri, M., et al. (2009). Feto-placental atherosclerotic lesions in intrauterine fetal demise: role of parental cigarette smoking. *Open Cardiovasc Med J.* 3: 51-56.
- Meigs, J.B., Shrader, P., Sullivan, L.M., et al. (2008). Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med.* 359(21): 2208-2219.

- Mendonça, I., dos Reis, R.P., Pereira, A., et al. (2011). Independent association of the variant rs1333049 at the 9p21 locus and coronary heart disease. *Rev Port Cardiol.* 30(6): 575-591.
- Moore, K.J., Tabas, I. (2011). Macrophages in the pathogenesis of atherosclerosis. *Cell.* 145(3): 341-355.
- Muendlein, A., Geller-Rhomberg, S., Saely, C.H., et al. (2009). Significant impact of chromosomal locus 1p13.3 on serum LDL cholesterol and on angiographically characterized coronary atherosclerosis. *Atherosclerosis.* 206(2): 494-499.
- Musunuru, K., Kathiresan, S. (2010). Genetics of coronary artery disease. *Annu Rev Genom Hum G.* 11: 91-108.
- Musunuru, K., Strong, A., Frank-Kamenetsky, M., et al. (2010). From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature.* 466(7307): 714-719.
- Myocardial Infarction Genetics Consortium, Kathiresan, S., Voight, B.F., et al. (2009). Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet.* 41(3): 334-341.
- Nichols, M., Townsend, N., Scarborough, P., et al. (2012). European Cardiovascular Disease Statistics 2012 Edition. Department of Public Health, University of Oxford. European Heart Network and European Society of Cardiology.
- Nielsen, L.B., Juul, K., Nordestgaard, B.G. (1998). Increased degradation of lipoprotein(a) in atherosclerotic compared with nonlesioned aortic intima-inner media of rabbits: in vivo evidence that lipoprotein(a) may contribute to foam cell formation. *Arterioscler Thromb Vasc Biol.* 18(4): 641-649.
- Nielsen, M.S., Madsen, P., Christensen, E.I., et al. (2001). The sortilin cytoplasmic tail conveys Golgi-endosome transport and binds the VHS domain of the GGA2 sorting protein. *EMBO J.* 20(9): 2180-2190.
- Nucifora, G., Schuijf, J.D., van Werkhoven, J.M., et al. (2009). Prevalence of coronary artery disease across the Framingham risk categories: coronary artery calcium scoring and MSCT coronary angiography. *J Nucl Cardiol.* 16(3): 368-375.
- Öörni, K., Pentikainen, M.O., Ala-Korpela, M., et al. (2000). Aggregation, fusion, and vesicle formation of modified low density lipoprotein particles: molecular mechanisms and effects on matrix interactions. *J Lipid Res.* 41(11): 1703-1714.
- Otsuka, T., Kawada, T., Ibuki, C., et al. (2010). Association between high-sensitivity cardiac troponin T levels and the predicted cardiovascular risk in middle-aged men without overt cardiovascular disease. *Am Heart J.* 159(6): 972-978.
- Paré, G., Chasman, D.I., Kellogg, M., et al. (2008). Novel association of ABO histo-blood group antigen with soluble ICAM-1: results of a genome-wide association study of 6,578 women. *PLoS Genet.* 4(7): e1000118.
- Paterson, A.D., Lopes-Virella, M.F., Waggott, D., et al. (2009). Genome-wide association identifies the ABO blood group as a major locus associated with serum levels of soluble E-selectin. *Arterioscler Thromb Vasc Biol.* 29(11): 1958-1967.
- Paynter, N.P., Chasman, D.I., Paré, G., et al. (2010). Association between a literature-based genetic risk score and cardiovascular events in women. *JAMA.* 303(7): 631-637.
- Peden, J.F., Farrall, M. (2011). Thirty-five common variants for coronary artery disease: the fruits of much collaborative labour. *Hum Mol Genet.* 20(R2): R198-205.
- Pericak-Vance, M.A. (2001). Analysis of genetic linkage data for Mendelian traits. *Curr Protoc Hum Genet.* Chapter 1: Unit 1.4.
- Poon, M., Zhang, X., Dunskey, K.G., et al. (1997). Apolipoprotein(a) induces monocyte chemotactic activity in human vascular endothelial cells. *Circulation.* 96(8): 2514-2519.
- Pranavchand, R., Reddy, B.M. (2013). Current status of understanding of the genetic etiology of coronary heart disease. *J Postgrad Med.* 59(1): 30-41.
- Psaty, B.M., Anderson, M., Kronmal, R.A., et al. (2004). The association between lipid levels

- and the risks of incident myocardial infarction, stroke, and total mortality: the cardiovascular health study. *J Am Geriatr Soc.* 52(10): 1639-1647.
- Pu, X., Xiao, Q., Kiechl, S., et al. (2013). ADAMTS7 cleavage and vascular smooth muscle cell migration is affected by a coronary-artery-disease-associated variant. *Am J Hum Genet.* 92(3): 366-374.
- Rader, D.J., Cohen, J., Hobbs, H.H. (2003). Monogenic hypercholesterolemia: new insights in pathogenesis and treatment. *J Clin Invest.* 111(12): 1795-1803.
- Reilly, M.P., Li, M., He, J., et al. (2011). Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. *Lancet.* 377(9763): 383-392.
- Ridker, P.M., Buring, J.E., Rifai, N., et al. (2007). Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. *JAMA.* 297(6): 611-619.
- Ridker, P.M., Paynter, N.P., Rifai, N., et al. (2008). C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds risk score for men. *Circulation.* 118(22): 2243-2251.
- Ripatti, S., Tikkanen, E., Orho-Melander, M., et al. (2010). A multilocus genetic risk score for coronary heart disease: case-control and prospective cohort analyses. *Lancet.* 376(9750): 1393-1400.
- Risch, N., Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science.* 273(5281): 1516-1517.
- Roberts, R., Stewart, A.F. (2012a). Genes and Coronary Artery Disease – Where Are We? *J Am Coll Cardiol.* 60(18): 1715-1721.
- Roberts, R., Stewart, A.F. (2012b). 9p21 and the genetic revolution for coronary artery disease. *Clin Chem.* 58(1): 104-112.
- Ross, R. (1999). Atherosclerosis – an inflammatory disease. *N Engl J Med.* 340(2): 115-126.
- Rubinstein, A., Izhakov, E. (2011). Lipoprotein associated phospholipase A2. *Harefuah.* 150(2): 136-140.
- Saleheen, D., Alexander, M., Rasheed, A., et al. (2010). Association of the 9p21.3 locus with risk of first-ever myocardial infarction in Pakistanis: case-control study in South Asia and updated meta-analysis of Europeans. *Arterioscler Thromb Vasc Biol.* 30(7): 1467-1473.
- Samani, N.J., Braund, P.S., Erdmann, J., et al. (2008). The novel genetic variant predisposing to coronary artery disease in the region of the PSRC1 and CELSR2 genes on chromosome 1 associates with serum cholesterol. *J Mol Med (Berl).* 86(11): 1233-1241.
- Samani, N.J., Erdmann, J., Hall, A.S., et al. (2007). Genomewide association analysis of coronary artery disease. *N Engl J Med.* 357(5): 443-453.
- Sayols-Baixeras, S., Lluís-Ganella, C., Lucas, G., et al. (2014). Pathogenesis of coronary artery disease: focus on genetic risk factors and identification of genetic variants. *Appl Clin Genet.* 7: 15-32.
- Schunkert, H., König, I.R., Kathiresan, S., et al. (2011). Large-scale association analyses identifies 13 new susceptibility loci for coronary artery disease. *Nat Genet.* 43(4): 333-338.
- Shanker, J., Arvind, P., Jambunathan, S., et al. (2014). Genetic analysis of the 9p21.3 CAD risk locus in Asian Indians. *Thromb Haemost.* 111(5): 960-969.
- Shea, J., Agarwala, V., Philippakis, A.A., et al. (2011). Comparing strategies to fine-map the association of common SNPs at chromosome 9p21 with type 2 diabetes and myocardial infarction. *Nat Genet.* 43(8): 801-805.
- Shen, G.Q., Li, L., Rao, S., et al. (2008). Four SNPs on chromosome 9p21 in a South Korean population implicate a genetic locus that confers high cross-race risk for development of

- coronary artery disease. *Arterioscler Thromb Vasc Biol.* 28(2): 360-365.
- Southam, L., Panoutsopoulou, K., Rayner, N.W., et al. (2011). The effect of genome-wide association scan quality control on imputation outcome for common variants. *Eur J Human Genet.* 19(5): 610–614.
- Stranger, B.E., Stahl, E.A., Raj, T. (2011). Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics.* 187(2): 367-383.
- Suchindran, S., Rivedal, D., Guyton, J.R., et al. (2010). Genome-wide association study of Lp-PLA(2) activity and mass in the Framingham Heart Study. *PLoS Genet.* 6(4): e1000928.
- Swerdlow, D.I., Holmes, M.V., Harrison, S., et al. (2012). The genetics of coronary heart disease. *Br Med Bull.* 102(1): 59-77.
- Tabas, I. (2010). Macrophage death and defective inflammation resolution in atherosclerosis. *Nat Rev Immunol.* 10(1): 36-46.
- Tabas, I., Williams, K.J., Boren, J. (2007). Subendothelial lipoprotein retention as the initiating process in atherosclerosis: update and therapeutic implications. *Circulation.* 116(16): 1832-1844.
- Tall, A.R., Ai, D. (2011). Sorting out sortilin. *Circ Res.* 108(2): 158-60.
- Teslovich, T.M., Musunuru, K., Smith, A.V., et al. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. *Nature.* 466(7307): 707-713.
- Thanassoulis, G., Vasan, R.S. (2010). Genetic cardiovascular risk prediction – will we get there? *Circulation.* 122(22): 2323-2334.
- Trégouët, D.A., König, I.R., Erdmann, J., et al. (2009). Genome-wide haplotype association study identifies the SLC22A3-LPAL2-LPA gene cluster as a risk locus for coronary artery disease. *Nat Genet.* 41: 283-285.
- Vargas-Alarcón, G., Posados-Romero, C., Villarreal-Molina, T., et al. (2013). Single nucleotide polymorphisms within LIPA (Lysosomal Acid Lipase A) gene are associated with susceptibility to premature coronary artery disease. A replication in the Genetic of Atherosclerotic Disease (GEA) Mexican study. *PLoS One.* 8(9): e74703.
- Varret, M., Rabès, J.P., Saint-Jore, B., et al. (1999). A third major locus for autosomal dominant hypercholesterolemia maps to 1p34.1-p32. *Am J Hum Genet.* 64(5): 1378-1387.
- Vincent, J.B., Skaug, J., Scherer, S.W. (2000). The human homologue of flamingo, EGFL2, encodes a brain-expressed large cadherin-like protein with epidermal growth factor-like domains, and maps to chromosome 1p13.3-p21.1. *DNA Res.* 7(3): 233-235.
- Visscher, P.M., Brown, M.A., McCarthy, M.I., et al. (2012). Five years of GWAS discovery. *Am J Hum Genet.* 90(1): 7-24.
- Wang, L., Fan, C., Topol, S.E., et al. (2003). Mutation of MEF2A in an inherited disorder with features of coronary artery disease. *Science.* 302(5650): 1578-1581.
- Wang, L., Zheng, J., Bai, X., et al. (2009). ADAMTS-7 mediates vascular smooth muscle cell migration and neointima formation in balloon-injured rat arteries. *Circ Res.* 104(5): 688-698.
- Wang, Q. (2005). Molecular genetics of coronary artery disease. *Curr Opin Cardiol.* 20(3): 182-188.
- Wellcome Trust Case Control Consortium. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 447(7145): 661-678.
- Wild, P.S., Zeller, T., Schillert, A., et al. (2011). A genome-wide association study identifies LIPA as a susceptibility gene for coronary artery disease. *Circ Cardiovasc Genet.* 4(4): 403-412.
- Willer, C.J., Sanna, S., Jackson, A.U., et al. (2008). Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet.* 40(2): 161-169.

- Willyard, C. (2013). Pathology: at the heart of the problem. *Nature*. 493: S10-11.
- Wilson, P.W., D'Agostino, R.B., Levy, D., et al. (1998). Prediction of coronary heart disease using risk factor categories. *Circulation*. 97: 1837-1847.
- Witztum, J.L., Lichtman, A.H. (2014). The influence of innate and adaptive immune responses on atherosclerosis. *Annu Rev Pathol*. 9: 73-102.
- Woodward, M., Brindle, P., Tunstall-Pedoe, H. (2007). Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). *Heart*. 93(2): 172-176.
- Yoshikawa, Y., Satoh, T., Tamura, T., et al. (2007). The M-Ras-RA-GEF-2-Rap1 pathway mediates tumor necrosis factor-alpha dependent regulation of integrin activation in splenocytes. *Mol Biol Cell*. 18(8): 2949-2959.
- Zeller, T., Blankenberg, S., Diemert, P. (2012). Genomewide association studies in cardiovascular disease – an update 2011. *Clin Chem*. 58(1): 92-103.
- Zethelius, B., Johnston, N., Venge, P. (2006). Troponin I as a predictor of coronary heart disease and mortality in 70-year-old men: a community-based cohort study. *Circulation*. 113(8): 1071-1078.
- Zheng, G.H., Chen, H.Y., Xiong, S.Q., et al. (2011). Lipoprotein-associated phospholipase A2 gene V279F polymorphisms and coronary heart disease: a meta-analysis. *Mol Biol Rep*. 38(6): 4089-4099.
- Zhu, M., Zhao, S. (2007). Candidate gene identification approach: progress and challenges. *Int J Biol Sci*. 3(7): 420-427.
- Zschenker, O., Illies, T., Ameis, D. (2006). Overexpression of lysosomal acid lipase and other proteins in atherosclerosis. *J Biochem*. 140(1): 23-38.

## **KASUTATUD VEEBIAADRESSID**

<http://rhk.sm.ee/>  
[http://www.epi.bris.ac.uk/CVDethrisk/CHD\\_CVD\\_form.html](http://www.epi.bris.ac.uk/CVDethrisk/CHD_CVD_form.html)  
<http://www.genome.gov/gwastudies/>  
<http://www.ncbi.nlm.nih.gov/SNP/>  
[http://www.who.int/cardiovascular\\_diseases/publications/atlas\\_cvd/en/](http://www.who.int/cardiovascular_diseases/publications/atlas_cvd/en/)  
<https://www.broadinstitute.org/mpg/snap/ldsearch.php>  
[www.R-project.org](http://www.R-project.org)



# LISA

**Tabel 3.** Lisaandmed riskiskooride arvutamisel kasutatud 33 markeri kohta.

SNP-i ID (riskialleel)	Imputeeritud/ genotüpiseeritud	Imputeerimis- kvaliteet	Proxy-markeri ID (riskialleel)
rs1333049 (C)	Imputeeritud	0,918	rs4977574 (G)
rs602633 (C)	Imputeeritud	0,968	rs599839 (A)
rs17114036 (A)	Imputeeritud	0,995	-
rs11556924 (C)	Genotüpiseeritud		
rs964184 (G)	Imputeeritud	0,998	rs17120029 (T)
rs579459 (C)	Genotüpiseeritud		
rs1412444 (T)	Genotüpiseeritud		
rs7173743 (T)	Genotüpiseeritud		
rs9818870 (T)	Genotüpiseeritud		
rs12190287 (C)	Imputeeritud	0,986	-
rs9515203 (T)	Genotüpiseeritud		
rs3798220 (C)	Imputeeritud	0,903	-
rs3184504 (T)	Genotüpiseeritud		
rs2895811 (C)	Imputeeritud	0,997	rs17562391 (T)
rs1746048 (C)	Imputeeritud	0,999	rs607363 (A)
rs9982601 (T)	Imputeeritud	0,994	-
rs12936587 (G)	Imputeeritud	0,993	rs12449964 (C)
rs7025486 (A)	Imputeeritud	0,937	-
rs1122608 (G)	Imputeeritud	0,982	rs3786725 (G)
rs12413409 (G)	Imputeeritud	0,999	rs12221064 (C)
rs12526453 (C)	Imputeeritud	0,997	rs7739181 (G)
rs216172 (C)	Imputeeritud	0,997	rs11655813 (T)
rs6725887 (C)	Imputeeritud	0,999	rs7582720 (C)
rs3869109 (G)	Genotüpiseeritud		

rs974819 (T)	Genotüpeeritud		
rs4773144 (G)	Genotüpeeritud		
rs17465637 (C)	Genotüpeeritud		
rs17609940 (G)	Imputeeritud	0,998	-
rs46522 (T)	Imputeeritud	1,000	rs2546491 (G)
rs11206510 (T)	Genotüpeeritud		
rs10953541 (C)	Imputeeritud	0,999	rs7785962 (G)
rs2075650 (G)	Genotüpeeritud		
rs2505083 (C)	Imputeeritud	0,909	-

*Proxy*-marker – imputeeritud markerile valitud asendusmarker tingimusel LD=1, mis on Illumina HumanOmniExpress kiibiga genotüpeeritud; - – puudub *proxy*-marker tingimusel LD=1.

# LIHTLITSENTS

Mina, Anett Tähiste

(sünnikuupäev: 04.11.1991)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose

Geneetilistel markeritel põhinev südame isheemiatõve riski hindamine  
Eesti Geenivaramu kohordis,

mille juhendaja on Maris Alver

- 1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
- 1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.
3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 26.05.2014